THE SECOND ASIAN CONFERENCE ON INFORMATION SYSTEMS

The proceedings

ACIS 2013

Phuket, Thailand

October 31 – November 2, 2013
Phuket, Thailand

Proceedings of

Asian Conference on Information Systems 2013

ACIS 2013

October 31 – November 2, 2013
Phuket Thailand

Organizers:

Sirindhorn International Institute of Technology (SIIT)

Thammasat University

Prince of Songkla University (PSU)

Sponsors:

Thailand Convention & Exhibition Bureau (TCEB)

Panyapiwat Institute of Management (PIM)

Rangsit University (RSU)

Collaborative Institutes:

The Institute of Electrical Engineers of Japan (IEEJ)

The International Institute of Applied Informatics, Japan (IIAI)

Japan Advanced Institute of Science and Technology, Japan (JAIST)

National Electronics and Computer Technology Center

About this publication

Title:	:	Proceedings of the Second Asian Conference on Information Systems (ACIS 2013)
Program Chairs	:	Kiyota Hashimoto (Osaka Prefecture University, Japan) Thanaruk Theeramunkong (SIIT, Thammasat University, Thailand)
Program Co-Chairs	:	Hidekazu Yanagimoto (Osaka Prefecture University, Japan) Thepchai Supnithi (NECTEC, Thailand)
Secretary General	:	Waree Kongprawechnon (SIIT, Thammasat University, Thailand)
Production Assistant	:	Eakasit Pacharawongsakda (SIIT, Thammasat University, Thailand) Nattapong Tongtep (SIIT, Thammasat University, Thailand)
Cover Designer	:	Kamonnut Bangwaek (SIIT, Thammasat University, Thailand)
Printing Production	:	Sirindhorn International Institute of Technology, Thammasat University 131 Moo 5 Tiwanont Road, Bangkadi, Muang Pathumthani 12000 Thailand Tel.: +66-2-501-3505(-20) Fax.: +66-2-501-3524
Date Published	:	October 2013
Organizers	:	Sirindhorn International Institute of Technology Thammasat University Osaka Prefecture University (OPU, Japan) Japan Advanced Institute of Science and Technology (JAIST, Japan) National Electronics and Computer Technology Center (NECTEC, Thailand) Panyapiwat Institute of Management (PIM, Thailand) Rangsit University (RSU, Thailand) Prince of Songkla University (PSU, Thailand)
Sponsors	:	Thailand Convention & Exhibition Bureau (TCEB) Panyapiwat Institute of Management (PIM) Rangsit University (RSU)
Contact	:	Sirindhorn International Institute of Technology, Thammasat University 131 Moo 5 Tiwanont Road, Bangkadi, Muang Pathumthani 12000 Thailand Tel.: +66-2-501-3505(-20) Fax.: +66-2-501-3524 e-mail: thanaruk@siit.tu.ac.th URL: saki.siit.tu.ac.th/acis2013

© 2013 by ACIS 2013

670 p. 1.71 cm. – (2013)
Includes bibliographical references.
ISBN: 978-616-91886-0-5
Printed in Thailand
05 06 07 08 09 5 4 3 2 1

Message from General Chairs

As the General Chair, I would like to congratulate on the success of the wonderful arrangement of the Second Asian Conference on Information Systems (ACIS-2013) held in Phuket, Thailand. The Second Asian Conference on Information Systems, ACIS 2013, following the successes of four Japan-China Joint Symposia of Information Systems (JCIS 2008-2011), the Japan-Cambodia Joint Symposium of Information Systems and Communication Technology (JCAICT 2011) and ACIS2012 (the first conference of the ACIS series) is the unique, premier meeting of researchers, professionals and practitioners to discuss various current issues in information systems and related areas with a special focus on the current conditions and the future sustainable developments of Asian countries. The second conference is hosted by Sirindhorn International Institute of Technology, Thammasat University (SIIT), National Electronics and Computer Technology Center (NECTEC), Panyapiwat Institute of Management (PIM), Rangsit University (RSU) and Prince of Songkla University (PSU). The collaborative institutes are the Institute of Electrical Engineers of Japan (IEEJ), the International Institute of Applied Informatics, Japan (IIAI), Osaka Prefecture University (OPU, Japan), and Japan Advanced Institute of Science and Technology (JAIST, Japan). The conference was partially sponsored by Thailand Convention & Exhibition Bureau (TCEB), Panyapiwat Institute of Management (PIM) and Rangsit University (RSU). This event aims to bring together many scholars, researchers and managers of various areas and industries for intellectual exchanges, research cooperation and professional development in the fields of information technology, information science, information management, computer science, knowledge science, knowledge management and service science. We are pleased to have attracted 106 excellent papers from all over the Asian countries, including Thailand, Japan, China, Korea, Vietnam, Indonesia, etc. In addition, we also have a prominent keynote speaker and two invited speakers that will clearly add a high note to this important event. Taking this opportunity, we would like to thank all the presenters that have contributed to this conference. Many thanks go to all the members for their dedication and hard work to make this a truly successful international conference. We sincerely hope that all of you enjoy this remarkable event.

October 2013
The ACIS 2013 General Chair:

Youji Kohda	(JAIST, Japan)
Pansak Siriruchatapong	(NECTEC, Thailand)

Message from Program Chairs

The Second Asian Conference on Information Systems (ACIS 2013) is held in Phuket, Thailand, during October 31 - November 2, 2013, follows its first meeting in Siem Reap, Cambodia in December 6-8, 2012. ACIS 2013 is a newly established international conference for the broad areas of information systems to provide an international forum for researchers and industry practitioners to share their new ideas, original research results and practical development experiences. The topics include artificial intelligence, data/text mining, communication technology, image processing, pattern recognition, information security, intelligent system development, ubiquitous computing, RFID technology, e-health, e-education, e-government, ICT For Development (ICT4D/ICTD), knowledge-based systems, software engineering, learning-supported systems, innovative business models, service science, management, and engineering, knowledge management and engineering, and logistics management and innovation. ACIS 2013 is hosted by Sirindhorn International Institute of Technology, Thammasat University (SIIT), National Electronics and Computer Technology Center (NECTEC), Panyapiwat Institute of Management (PIM), Rangsit University (RSU) and Prince of Songkla University (PSU). The collaborative institutes are the Institute of Electrical Engineers of Japan (IEEJ), the International Institute of Applied Informatics, Japan (IIAI), Osaka Prefecture University (OPU, Japan), and Japan Advanced Institute of Science and Technology (JAIST, Japan). The conference is partially sponsored by Thammasat University, Thailand Convention & Exhibition Bureau (TCEB), Panyapiwat Institute of Management (PIM) and Rangsit University (RSU).

In 2013, we received 131 research papers from various countries and regions in Asia, Australia, North America, South America, Europe, and Africa. Every submission was rigorously reviewed by at least two reviewers. The initial results were discussed among the reviewers and finally judged by the Program Committee Chairs. When there was a conflict on the result, an additional review was provided by Program Committee Chairs. The program committee members were deeply involved in a highly selective process. As a result, only 43 papers (approximately 32.82% of the 131 submitted papers) were accepted as regular papers, 30 papers (22.90% of them) as posters, and 33 papers (25.19% of them) as special session papers.

The ACIS 2013 conference program included four special sessions: Special Session on Service Sciences (ACIS-SS), Special Session on Computational Medical Image Processing (ACIS-CMIP), Special Session on Onomatopoeia Human Computer Interaction (ACIS-OHCI), and Special Session on Information Technology Management (ACIS-ITM). ACIS 2013 would not have been successful without the support of committee members, reviewers, workshop organizers, tutorial speakers, and invited speakers, and organizing staffs.

We are indebted to the members of the Steering Committee for their invaluable suggestions and supports throughout the organization process. We highly appreciate the Program Committee members, and external reviewers for their technical effort in providing straightforward scientific comments and impartial judgments in the review process of ACIS 2013. We wish to thank the Special Session Chairs; Michitaka Kosaka (JAIST, Japan), Stanislav Makhanov (SIIT, Thailand), M.L. Kunthorn Kasemsan (Rangsit University, Thailand), and Takanori Komatsu and Satoshi Nakamura (Meiji University, Japan) for their excellent arrangement of the Special Sessions. Many thanks are given to the distinguished keynote speakers, invited speakers; Riichiro Mizoguchi (JAIST), Patrick Doherty (Linkopings Universitet) and Siriwan Suebnukarn (Thammasat University) for their attractive and motivational talks and lectures. We thank the Honorary Co-Chair, Norihisa Komoda (Osaka University, Japan) and Somnuk Tangtermsirikul (SIIT, Thailand) as well as the General Co-Chair, Yoji Kohda (JAIST, Japan), Pansak Siriruchatapong (NECTEC, Thailand) for their kind support and their useful guidance and their sharp advice in various spectrums related to conference arrangement. We are also grateful to the Secretary General, Waree Kongprawechnon and our Local Arrangement Committee, including Boontawee

Suntisrivaraporn, Pakinee Aimmanee, Suthathip Suanmali and Sasiporn Usanavasin from SIIT (Thammasat University), Somrote Komolavanij, Datchakorn Tancharoen and Ruengsak Kawtummachai from PIM, Todsanai Chumwatana and Tanatorn Tanantong from RSU, Rattana Wetprasit from PSU, Thatsanee Charoenporn and Ithipan Methasate from NECTEC. We are also indebted to Eakasit Pacharawongsakda and Kamonnut Bangwaek for their great effort in preparing the program book and proceedings of this conference. Thanks to our strong staffs; Choermath Hongakkarapun, Cholthicha Pradidkwan, Nongnuch Ketui, Nattapong Tongtep, Tanatorn Tanantong, Nichnan Kittiphattanabawon, Natthapon Pannurat, Sawit Kasuriya, Krich Nasingkun, Nattapol Kritsuthikul, Suthum Keerativittayanun, Tama Duangnamol, Nanthawadee Sucharittham, Siriwon Taewijit, Pasinee Apisakmontri, Thanasarn Ternhernhong and Wirat Chinnan for their unlimited help towards the success of the conference. Also thanks to the Session Chairs; Weerawut Thanhikam, Sasiporn Usanavasin, Todsanai Chumwatana, Tu Bao Ho, Poonpong Boonbrahm, Somchai Lekchaleon, Tomoharu Nakashima, Takuya Maekawa, Michitaka Kosaka, Naoki Mori, Stanislav S. Makhanov and Takanori Komatsu. While the arrangement of ACIS 2013 involved so many people, we would like to make an additional thank to the contributors who gave their hand for ACIS 2013 but their names may not be listed.

Finally we wish to thank all authors and all conference participants for their contribution and support. We hope all participants took this opportunity to share and exchange ideas with the others and enjoyed ACIS 2013 in the wonderful city - Phuket, Thailand.

October 2013

Thanaruk Theeramunkong	(SIIT, Thammasat University, Thailand)
Kiyota Hashimoto	(Osaka Prefecture University, Japan)
Hidekazu Yanagimoto	(Osaka Prefecture University, Japan)
Thepchai Supnithi	(NECTEC, Thailand)

Program Chairs and Vice Chairs
ACIS 2013

Message from Local Arrangement Chairs

We are pleased to invite you to join the Second Asian Conference on Information Systems (ACIS 2013), held this year in Phuket, Thailand, during October 31 - November 2, 2013. The conference is arranged under the collaboration between Sirindhorn International Institute of Technology, Thammasat University (SIIT), National Electronics and Computer Technology Center (NECTEC), Panyapiwat Institute of Management (PIM), Rangsit University (RSU) and Prince of Songkla University (PSU). To arrange the conference, we are indebted to several members at PSU, NECTEC, SIIT, PIM and RSU, including their faculty members, staffs and students. As Thailand's largest island with the size close to Singapore, Phuket is a well-known business and sightseeing province in the southern of Thailand, connected to mainland Thailand by two bridges. Situated off the west coast of Thailand in the Andaman Sea, Phuket obtains its wealth from tin and rubber, and enjoys a rich and colorful history. The island was on one of the major trading routes between India and China, and was frequently mentioned in foreign ship logs of Portuguese, French, Dutch and English traders. The region now derives much of its income from tourism.

With great honor and pleasure, we are glad to make local arrangement of the ACIS 2013 in Phuket and welcome the approximately 140 participants, including the committee members, invited speakers, authors, presenters and audiences. Finally we wish to thank all committee members for their great effort, contribution and support to establish this precious academic conference in Phuket. We hope all participants take this opportunity to share and exchange ideas with the others and enjoyed ACIS 2013 and enjoy their stay in our wonderful Phuket.

October 2013

ACIS 2013 Local Arrangement Representatives:

Rattana Wetprasit	(Prince of Songkla University, Thailand)
Veeraporn Siddoo	(Prince of Songkla University, Thailand)
Panisa Treepong	(Prince of Songkla University, Thailand)
Kitsiri Chochiang	(Prince of Songkla University, Thailand)

ACIS 2013 Arrangement Committee members:

Boontawee Suntisrivaraporn	(SIIT, Thammasat University, Thailand)
Chawalit Jeenanunta	(SIIT, Thammasat University, Thailand)
Pakinee Aimmanee	(SIIT, Thammasat University, Thailand)
Sasiporn Usanavasin	(SIIT, Thammasat University, Thailand)
Suebsak Nanthavanij	(SIIT, Thammasat University, Thailand)
Suthathip Suanmali	(SIIT, Thammasat University, Thailand)
Waree Kongprawechnon	(SIIT, Thammasat University, Thailand)
Somrote Komolavanij	(Panyapiwat Institute of Management, Thailand)
Datchakorn Tancharoen	(Panyapiwat Institute of Management, Thailand)
Ruengsak Kawtummachai	(Panyapiwat Institute of Management, Thailand)
Kulthon Kasemsan	(Rangsit University, Thailand)
Tanatorn Tanantong	(Rangsit University, Thailand)
Todsanai Chumwatana	(Rangsit University, Thailand)

ACIS 2013 Organizing Committees

Organizing Committee

Honorary Co-Chair:

Norihisa Komoda	(Osaka University, Japan)
Somnuk Tangtermsirikul	(SIIT, Thammasat University, Thailand)

General Co-Chair:

Yoji Kohda	(JAIST, Japan)
Pansak Siriruchatapong	(NECTEC, Thailand)

International Advisory Committee:

Norihisa Komoda	(Osaka University, Japan)
Michitaka Kosaka	(JAIST, Japan)
Hiroshi Tsuji	(Osaka Prefecture University, Japan)
Badaruddin Mohamed	(Universiti Sains Malaysia – USM, Malaysia)
Kulthon Kasemsan	(Rangsit University, Thailand)
Phal Des	(Royal University of Phnom Penh, Cambodia)
Andrzej Skulimowski	(AGH University of Science and Technology, Poland)

Program Co-Chair:

Kiyota Hashimoto	(Osaka Prefecture University, Japan)
Thanaruk Theeramunkong	(SIIT, Thammasat University, Thailand)

Vice Program Co-Chair:

Hidekazu Yanagimoto	(Osaka Prefecture University, Japan)
Thepchai Supnithi	(NECTEC, Thailand)

Publication Co-Chair:

Takuya Maekawa	(Osaka University, Japan)
Datchakorn Tancharoen	(Panyapiwat Institute of Management, Thailand)

Publicity Co-Chair:

Masaki Yumoto	(Kinki University, Japan)
Pisit Charnkeitkong	(Panyapiwat Institute of Management, Thailand)

Financial Co-Chair:

Yoji Taniguchi	(Hitachi, Japan)
Pakinee Aimmanee	(SIIT, Thammasat University, Thailand)

Secretary General:

Waree Kongprawechnon	(SIIT, Thammasat University, Thailand)

Local Arrangement Co-Chair:

Rattana Wetprasit	(Prince of Songkla University, Thailand)
Veeraporn Siddoo	(Prince of Songkla University, Thailand)
Panisa Treepong	(Prince of Songkla University, Thailand)
Kitsiri Chochiang	(Prince of Songkla University, Thailand)
Somrote Komolavanij	(Panyapiwat Institute of Management, Thailand)
Chawalit Jeenanunta	(SIIT, Thammasat University, Thailand)
Suebsak Nanthavanij	(SIIT, Thammasat University, Thailand)

Local Arrangement Members:

Boontawee Suntisrivaraporn	(SIIT, Thammasat University, Thailand)
Sasiporn Usanavasin	(SIIT, Thammasat University, Thailand)
Suthathip Suanmali	(SIIT, Thammasat University, Thailand)
Chutima Beokhaimook	(Rangsit University, Thailand)
Todsanai Chumwatana	(Rangsit University, Thailand)
Ruengsak Kawtummachai	(Panyapiwat Institute of Management, Thailand)

Secretariat:

Thatsanee Charoenporn	(NECTEC, Thailand)
Choermath Hongakkarapun	(SIIT, Thammasat University, Thailand)
Cholthicha Pradidkwan	(SIIT, Thammasat University, Thailand)
Eakasit Pacharawongsakda	(SIIT, Thammasat University, Thailand)
Ithipan Methasate	(NECTEC, Thailand)
Kamonnut Bangwaek	(SIIT, Thammasat University, Thailand)
Krich Nasingkun	(SIIT, Thammasat University, Thailand)
Nanthawadee Sucharittham	(SIIT, Thammasat University, Thailand)
Nattapol Kritsuthikul	(SIIT, Thammasat University, Thailand)
Nattapong Tongtep	(SIIT, Thammasat University, Thailand)
Natthapon Pannurat	(SIIT, Thammasat University, Thailand)
Nichcha Puakanokhiran	(SIIT, Thammasat University, Thailand)
Nichnan Kittiphattanabawon	(Walailak University, Thailand)
Nongnuch Ketui	(SIIT, Thammasat University, Thailand)
Pairunya Chaichana	(SIIT, Thammasat University, Thailand)
Pasinee Apisakmontri	(SIIT, Thammasat University, Thailand)
Peechalika Sertsirikul	(SIIT, Thammasat University, Thailand)
Sawit Kasuriya	(SIIT, Thammasat University, Thailand)
Sirinart Sirijorn	(SIIT, Thammasat University, Thailand)
Siriwon Taewijit	(SIIT, Thammasat University, Thailand)
Suthum Keerativittayanun	(SIIT, Thammasat University, Thailand)
Tama Duangnamol	(SIIT, Thammasat University, Thailand)
Tanatorn Tanantong	(SIIT, Thammasat University, Thailand)
Thanasarn Ternhernhong	(SIIT, Thammasat University, Thailand)
Wirat Chinnan	(SIIT, Thammasat University, Thailand)
Witchapon Jomprapan	(SIIT, Thammasat University, Thailand)

Hosts and Co-Hosts:

Sirindhorn International Institute of Technology	(SIIT, Thammasat University, Thailand)
Japan Advanced Institute of Science and Technology	(JAIST, Japan)
National Electronics and Computer Technology Center	(NECTEC, Thailand)
Osaka Prefecture University	(OPU, Japan)
Panyapiwat Institute of Management	(PIM, Thailand)
Prince of Songkla University	(PSU, Thailand)
Rangsit University	(RSU, Thailand)

Sponsors:

National Electronics and Computer Technology Center	(NECTEC, Thailand)
Panyapiwat Institute of Management	(PIM, Thailand)
Rangsit University	(RSU, Thailand)
Thammasat University	(TU, Thailand)
Thailand Convention & Exhibition Bureau	(TCEB, Thailand)

List of Reviewers

Akara Sopharak	(Burapha University, Thailand)
Alexander Schill	(TU Dresden, Germany)
Ayako Hiramatsu	(Osaka Sangyo University, Japan)
Boontawee Suntisrivaraporn	(SIIT, Thammasat University, Thailand)
Bunyarit Uyyanonvara	(SIIT, Thammasat University, Thailand)
Chai Wutiwiwatchai	(NECTEC, Thailand)
Chanjira Sinthanayothin	(NECTEC, Thailand)
Chawalit Jeenanunta	(SIIT, Thammasat University, Thailand)
Choochart Haruechaiyasak	(NECTEC, Thailand)
Datchakorn Tancharoen	(Panyapiwat Institute of Management, Thailand)
Dinh Chi Hieu	(Ministry of Information and Communications, Vietnam)
Dosam Hwang	(Yeungnam University, Korea)
Eisuke Ito	(Kyushu University, Japan)
Ekawit Nantajeewarawat	(SIIT, Thammasat University, Thailand)
Gun Srijuntongsiri	(SIIT, Thammasat University, Thailand)
Helena Grunfeld	(Victoria University, Melbourne)
Hidekazu Yanagimoto	(Osaka Prefecture University, Japan)
Hiroyuki Kojima	(Hiroshima Institute Technology, Japan)
Ithipan Methasate	(NECTEC, Thailand)
Kazuhiro Takeuchi	(Osaka Electro-Communication University, Japan)
Kimtho Po	(Institute of Technology of Cambodia, Cambodia)
Kiyota Hashimoto	(Osaka Prefecture University, Japan)
Kritsada Sriphaew	(Rangsit University, Thailand)
Makoto Okada	(Osaka Prefecture University, Japan)
Masaki Yumoto	(Kinki University, Japan)
Masao Izumi	(Osaka Prefecture University, Japan)
Michitaka Kosaka	(Japan Advanced Institute of Science and Technology, Japan)
Ming Zhou	(Microsoft Research, China)
Naoki Mori	(Osaka Prefecture University, Japan)
Nguyen Duy Hung	(SIIT, Thammasat University, Thailand)
Nuanwan Soonthornphisaj	(Kasetsart University, Thailand)
Pakinee Aimmanee	(SIIT, Thammasat University, Thailand)
Poonpong Boonbrahm	(Walailak University, Thailand)
Rachada Kongkachandra	(Thammasat University, Thailand)
Rattana Wetprasit	(Prince of Songkla University, Thailand)
Rei Itsuki	(Hiroshima International University, Japan)
Ruengsak Kawtummachai	(Panyapiwat Institute of Management, Thailand)
Ryan Ko	(University of Waikato, New Zealand)
Ryoju HAMADA	(Tohoku Univesity, Japan)
Ryosuke Saga	(Osaka Prefecture University, Japan)
Sachio Hirokawa	(Kyushu University, Japan)

Sanparith Marukatat	(NECTEC, Thailand)
Sasiporn Usanavasin	(SIIT, Thammasat University, Thailand)
Shimpei Matsumoto	(Hiroshima Institute of Technology, Japan)
Sodany Kiv	(Université Catholique de Louvain, Belgium)
Sopheak Cheang	(National ICT Development Authority, Cambodia)
Stanislav Makhanov	(SIIT, Thammasat University, Thailand)
Stephane Bressan	(National University of Singapore, Singapore)
Steven Gordon	(SIIT, Thammasat University, Thailand)
Suebsak Nanthavanij	(SIIT, Thammasat University, Thailand)
Sumali Conlon	(University of Mississippi, United States)
Surapa Thiemjarus	(NECTEC, Thailand)
Suthathip Suanmali	(SIIT, Thammasat University, Thailand)
Takashi Moriyasu	(Toshiba Solutions Corporation, Japan)
Tatsuya Nakae	(Hitachi, Japan)
Teerayut Horanont	(The University of Tokyo, Japan)
Tetsuo Tanaka	(Kanagawa Institute of Technology, Japan)
Thanaruk Theeramunkong	(SIIT, Thammasat University, Thailand)
Thepchai Supnithi	(NECTEC, Thailand)
Todsanai Chumwatana	(Rangsit University, Thailand)
Tomoharu Nakashima	(Osaka Prefecture University, Japan)
Tomomi Kaneko	(Hokkaido Automotive Engineering College, Japan)
Toshiyuki Moritsu	(Hitachi, Japan)
Toshiyuki Ono	(Hitachi, Japan)
Vasutan Tunbunheng	(Panyapiwat Institute of Management, Thailand)
Virach Sornlertlamvanich	(NECTEC, Thailand)
Vlado Stankovski	(University of Ljubljana, Slovenia)
Waree Kongprawechnon	(SIIT, Thammasat University, Thailand)
Yoji Taniguchi	(Hitachi, Japan)

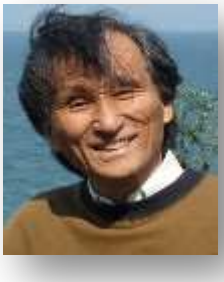
ACIS 2013

Asian Conference on Information Systems 2013

Keynote Speeches

Keynote Speech I

Ontological Engineering Approach to Real-World Problems



Prof. Riichiro Mizoguchi
Research Center for Service Science
Japan Advanced Institute of Science and Technology

There exist a lot of problems to be solved in the real world. People tend to solve them immediately after they realize needs to solve them. One of the issues here is that necessary consideration about the nature of those problems is often skipped to get solutions quickly, which sometimes leads to ad-hoc solutions and/or non-optimal solutions. Ontological engineering provides us with conceptual tools for analyzing problems in a right way by which we mean analysis of underlying background of the problems as well as their essential properties to obtain more general solutions. It also suggests that we should investigate the problems as deeply as possible like philosophers to reveal essential and intrinsic characteristics hidden in the superficial phenomena/appearance. Knowledge is necessarily something about existing entities and their relations, and ontology is an investigation of being, and hence ontology contributes to facilitation of our knowledge about the world in an essential manner. In my talk, after a brief introduction to ontological engineering, I explain how it works as a theory of content and/or content technology. It is amazing and surprising to learn how much we are ignorant of fundamental issues of the real world. Few people knows how process and event are different from each other, how property, attribute and quality are different, how many kinds of part-of relations exist, why is an inappropriate relation, why both <a piece of pie is-a pie> and <a piece of pie part-of a pie> are correct, etc. Ontological engineering answers you all such questions. After explaining theoretical issues of ontological engineering, I will talk about technological aspects of ontological engineering referring to my experiences. One of the important conceptual techniques is separation of what and how in procedures/algorithms. Then, I show you a couple of concrete examples of deployment of such conceptual tools in several domains.

Bibilography

Riichiro MIZOGUCHI received the B.S., M.S., and Ph.D. degrees from Osaka University, Osaka Japan, in 1972, 1974 and 1977, respectively. From 1978 to 1986 he was assistant professor in the Institute of Scientific and Industrial Research, Osaka University. From 1986 to 1989 he was associate professor, and professor from 1990 to 2012. He is currently professor of Research Center for Service Science, Japan Advanced Institute of Science and Technology. His research interests include knowledge-based systems, theories of role and function, ontological engineering, ontology development and intelligent learning support systems. He received honorable mention for the Pattern Recognition Society award, best paper award of the Institute of Electronics, Information and Communication Engineers award, 10th anniversary paper award from Japanese Society for Artificial Intelligence (JSAI), best paper award of ICCE99, best paper award of ICCE06, best paper award of JSAI, best paper award of JSiSE, best paper award of JSAI in 1985, 1988, 1996, 1999, 2006, 2006, 2010 and 2013, respectively. Dr. Mizoguchi was President of International AI in ED Society and Asia-Pacific Society for Computers in Education, Vice-president of Semantic Web Science Association (SWSA) and President of JSAI. He is currently associate editors of IEEE TLT and ACM TiiS.

Keynote Speech II

Collaborative Unmanned Aircraft Systems

Prof. Dr. Patrick Doherty
Dept. of Computer and Information Science,
Linköpings universitet

Automated specification, generation and execution of high level missions involving one or more heterogeneous unmanned aircraft systems is in its infancy. Much previous effort has been focused on the development of air vehicle platforms themselves together with the avionics and sensor subsystems that implement basic navigational skills. In order to increase the degree of autonomy in such systems so they can successfully participate in more complex mission scenarios such as those considered in emergency rescue that also include ongoing interactions with human operators, new architectural components and functionalities will be required to aid not only human operators in mission planning, but also the unmanned aircraft systems themselves in the automatic generation, execution and partial verification of mission plans to achieve mission goals. In this talk, I will present a delegation-based multi-agent framework that can be used for the automated planning, deployment and resource allocation of multiple unmanned aircraft systems in complex real-world missions. A prototype of the framework is operational in a number of autonomous unmanned aircraft developed in our research lab. Different scenarios involving UASs and emergency services assistance will be shown.

Bibilography

Patrick Doherty is a Professor of Computer Science at the Department of Computer and Information Sciences (IDA), Linköping University, Sweden. He heads the Artificial Intelligence and Integrated Computer Systems Division at IDA. He serves as director of LinkLab, a research center for future aviation systems, which is a collaborative endeavor between Linkoping University and Saab Aero Systems. He is an ECCAI fellow and currently an ECCAI board member. He has previously served as president of SAIS, the Swedish Artificial Intelligence Society. He is an associate editor for the Artificial Intelligence Journal and is on the board of directors for KR Inc. His main areas of interest are knowledge representation, automated planning, intelligent autonomous systems and multi-agent systems. A major area of application is with Unmanned Aircraft Systems.

His research group has successfully designed and deployed many autonomous UAV systems in the past decade have also won several international competitions pertaining to micro-aerial vehicles and to automated planning. He has over 150 refereed scientific publications in his areas of expertise and has given many keynote and invited talks at leading international conferences. He is also CEO of a new start-up company, UAS Technologies Sweden AB, which designs and markets Micro-Aerial Vehicles.

Keynote Speech III

Virtual Reality Dental Simulator

Assoc. Prof. Dr. Siriwan Suebnukarn
Advanced dental imaging and computing research center, Faculty of Dentistry
Thammasat University

Clinical training in developing nations faces a revolutionary opportunity for change created by the confluence of pressures from economic and societal changes along with opportunities for new modes of education offered by the advancement and widespread availability of Information and Communication Technology. Advances in learning technologies are creating opportunities to improve clinical training in 21st century competencies. In order to realize these benefits we need to move beyond models of content delivery and think afresh about the characteristics of stimulating and engaging learning environments. With recent advantages in biomedical imaging and virtual reality technology, training in dentistry is entering a time of change towards more realistic and interactive environments. I examine ways in which simulation environments and intelligent tutoring systems, can help to address some of the pedagogical challenges we face as educators. I illustrate with research studies from the work of my research group starting from an overview of haptic virtual reality dental simulator developed by our group. The results from clinical trials on pre-clinical training in endodontic access preparation and pre-surgical practice in endodontic microsurgery are presented and discussed. The strength of using a VR haptic system for clinical skill assessment is the ability to automatically record associated kinematic data which are not available in the conventional skill training environments. Such variables are needed to build algorithms for the new generation of improved clinical skill training system that may allow more effective training experience with real-time feedback of skill performance. The intelligent virtual reality dental training system could be an invaluable tool for training in developing nations where there may be a dearth of teaching dental practitioners and may point the way in the development of virtual reality training tools for other kinds of medical procedures.

Bibilography

Dr. Siriwan Suebnukarn, DDS, PhD is Associate Professor in the Faculty of Dentistry and Director of Advanced Dental Imaging and Computing Research Center at Thammasat University in Thailand. She has also served on the adjunct faculty at the Asian Institute of Technology. Dr. Suebnukarn's research has concentrated on the use of information and communication technology especially artificial intelligence to build intelligent clinical training and clinical decision support systems, and she has done pioneering work in the areas of intelligent tutoring and virtual reality simulation. Dr. Suebnukarn won the prestigious International Federation of Inventor Association's (IFIA) Lady Prize for the Best Women Invention in 2007, and the IFIA gold medals in World Cup for Computer Implemented Invention in 2008 and 2012 for her work on Virtual Reality Dental Simulator.

ACIS 2013

Asian Conference on Information Systems 2013

Regular Paper Session

CONTENTS

Page

Risk-aware Data Replication to Massively Multi-sites against Widespread Disasters 1
Shinya MATSUMOTO, Takaki Nakamura and Hiroaki MURAOKA

Evaluation of SSD Tier Method and SSD Cache Method in Tiered Storage System 8
Shinichi Hayashi and Norihisa KOMODA

Development and Experiment with Scheduling and Network Coding Tools in Wireless Testbed 15
Seng MOLIKA, Kong Marry and Phal DES

Noise Suppressor using Zero Phase Signal and Accumulated Spectrum Technique 21
Weerawut THANHIKAM

The Assignment of Chunk Size According to the Target Data Characteristics in Deduplication Backup System 29
Mikito Ogata and Norihisa KOMODA

Optimizing Acceleration-based Activity Models with User's Physical Characteristics Information 35
Yousuke Yasuda and Takuya MAEKAWA

Dynamic State Estimation for Vehicle Platoon System Based on Feedback Estimators 41
Hironori Suzuki and Takashi NAKATSUJI

New DEA Framework: Evaluating I/O Efficiency Structures 49
Kazushige INOUE, Kenji FUKUOKA, Shingo Aoki and Hiroshi TSUJI

Hybrid DEA Model Based on Hierarchical Approach 56
Kenji FUKUOKA, Yukie MAJIMA, Kazushige Inoue and Shingo AOKI

Markov Random Fields based Estimation of Distribution Algorithm and Application of Sharable Resource Constrained Scheduling Problem 62
X. HAO, J. TIAN, H. W. Lin and T. MURATA

New Best Solutions for the Close-Open Mixed Vehicle Routing Problem by using an Improved Harmony Search Algorithm 68
Tantikorn Pichpibul and Ruengsak KAWTUMMACHAI

Cognitive Technologies for Assessing Skills 73
Charles TIJUS, FranÃ§ois JOUEN, SÃ©bastien Poitrenaud and Michelle MOLINA

Planning Support in Capability Growth Trajectory Mining 81
Thongchai KAEWKIRIYA, Ryosuke Saga and Hiroshi TSUJI

Detecting Outlier in Graph Structure Data Using Centrality 88
Yukihiro TAKAYAMA, Ryosuke Saga and Takao MIYAMOTO

A Robust Privacy-Preserving Recommendation Algorithm 95
Alper BILGE, Ihsan Gunes and Huseyin POLAT

A Review on Data Mining Approach for Adverse Drug Reaction Research 103
Siriwon TAEWIJIT, Tu Bao Ho and Thanaruk THEERAMUNKONG

Data Mining for Life Insurance Knowledge Extraction: A Survey 111
Nanthawadee SUCHARITTHAM, Thanaruk THEERAMUNKONG, Choochart HARUECHAIYASAK, Bao Tu Ho and Dam Hieu CHI

Classification of Mammogram Images using Support Vector Machine 119
KARMILASARI, Suryarini WIDODO, Lussiana ETP, Matrissya Hermita and Lulu MAWADAH

Motion recognition of assistance for bed-wheelchair transferring from video 127
Masao IZUMI, Yurie Uno and Yukie MAJIMA

A Tennis Movie Retrieval System using Segmented-Swing Matching 133
Kaede YAMAZAKI, Masato Kasahara and Kosuke TAKANO

A Study on the Development of an Occupancy Simulator for Elderly Monitoring 141
Tomoharu NAKASHIMA, Yasuko MAEKAWA, Youhei KURODA, Ayaka Dobashi and Yumi HIGUCHI

Extracting Descriptive Life Profiles from Mobile GPS Data 147
Teerayut HORANONT, Apichon WITAYANGKURN, Yoshihide Sekimoto and Ryosuke SHIBASAKI

Towards the Use of Upper Ontologies for Refugee Emergencies in Disaster Management 153
Pasinee APISAKMONTRI, Ekawit NANTAJEEWARAWAT, Marut Buranarach and Mitsuru IKEDA

Line Segmentation for Khmer Printing Text Document 161
Dina SOK, Sovila SRUN, Por HENG, Beanbonyka RIM, Khoy Saovorak and Des PHAL

A Comparative Study on Confidence-Weighted Learning between Fuzzy and Non-Fuzzy Classifiers 169
Takeshi SUMITANI, Tomoharu Nakashima and Andrzej BARGIELA

Extending k-Representative Clustering Algorithm with an Information Theoretic-based Dissimilarity Measure for Categorical Objects 175
Thu-Hien Thi Nguyen and Van-Nam HUYNH

A Computational Method for Obtaining Stackelberg Solutions to Two-level Nonlinear Programming Problems through Strategic Oscillation Using Particle Swarm Optimization 182
Kosuke KATO, Shimpei MATSUMOTO, Shuhei MIMURA, Nobuyuki Yamawaki and Hideki KATAGIRI

Identifying Co-Creation Process in Energy Aggregation Services through Soft Systems Engineering 189
Motohisa FUNABASHI

An Enhancement of the Stakeholder Modeling Scheme for Requirements Engineering of Software Systems for Public 195
Jidtima SUNKHAMANI

Learning and Recognizing Behavior Patterns Based on the Distribution of the Direction of Movement 203
Kizuku NISHIMURA, Shigeki Aoki and Takao MIYAMOTO

Consideration of Page Dwell Time in B-to-B (Business-to-Business) Industry Web Analytics 210
Akiyuki Sekiguchi and Kazuhiko TSUDA

How Do Customers Buy Them at a Supermarket? Behavior Analysis from Real Observation and Agent Simulation 217
Masaki KITAZAWA, Fumiaki SATO, Takashi YAMADA, Masakazu Takahashi and Takao TERANO

Ontology-Enhanced Life Cycle Assessment: A Case Study of Application in Oil Refinery 223
Akkharawoot TAKHOM, Boontawee Suntisrivaraporn and Thepchai SUPNITHI

Content Management System for Food Industry 231
Tomoko Kashima and Shimpei MATSUMOTO

Engaging Reflective Thinking in Learning Mathematics Word Problems by Using a Personal Tutoring Agent 239
Tama DUANGNAMOL, Boontawee SUNTISRIVARAPORN, Thepchai Supnithi and Mitsuru IKEDA

A Virtual Environment for English as Foreign Language Learning Platform (veEFL): Applied "Single Idea of Concept" to Improve Writing skill of Low English Proficiency Students 247
Nattapol KRITSUTHIKUL, Shinobu HASEGAWA, , Cholwich Nattee and Thepchai SUPNITHI

An Image Creating System based on Kansei of Natural Language Sentences 255
Miki UENO, Akihiko YASUI, Naoki Mori and Keinosuke MATSUMOTO

A Method to Classify Sentences of Topics and Opinion in Online Short Letters about Stock Trading Information Using Support Vector Machine 263
Makoto OKADA, Kouhei Shimada and Kiyota HASHIMOTO

Topic Sentence Extraction from Editorial Articles Based on Sentence Structure and Topic Relevance 269
Muneaki HARADA, Hidekazu Yanagimoto and Michifumi YOSHIOKA

Exploiting WordNet for Wikipedia-Based Named Entity Disambiguation 275
Thuan K. Nguyen and Tru H. CAO

Evaluation of Web Pages Using Wikipedia 282
Shoki HIRATA, Hidekazu Yanagimoto and Michifumi YOSHIOKA

Evaluative Phrase Extraction from Customer Review 286
Satoshi SAKEMI, Hidekazu Yanagimoto and Michifumi YOSHIOKA

Using Core Phrases and Information Features for Keyphrase Extraction 291
Huyen K. Nguyen and Tru H. CAO

Robust Jointly Sparse Elastic Regression with Matrix Uncertainty 297
Seppo KARRILA

The Extracting of Shirt Composition using Canny Edge Detection with Close Contour 306
Lanjakorn Sewata and Salil BOONBRAHM

Digitally Generated Khmer Ornament for Khmer Architecture 312
Puthnith VAR, Sokha HENG, Por HENG, Saovorak KHOY, Pheakdey Nguonphan and Des PHAL

Patent Citation Recommendation Based on Topic Model Expansion 316
Supaporn Tantanasiriwong and Choochart HARUECHAIYASAK

Chinese and Japanese Word Learning System by Estimation of Word Difficulty 320
Peiqing ZANG, Hiromitsu SHIINA, Nobuyuki Kobayashi and Fumio KITAGAWA

A web based Programming Learning Management System with JavaScript Libraries 324
Shimpei MATSUMOTO, Shuichi Yamagishi and Tomoko KASHIMA

Estimating Degree of Concentration through Activity Recognition: Use in a Classroom with SNS 330
Tomoaki UEDA, Hiroshi SUGIMURA, Tetsuo Tanaka and Kazunori MATSUMOTO

IT Governance in Shared Services in Public Sector 332
Omar M. ALI

e-Government and its Implementation Challenges in Nepal 340
Subarna SHAKYA

ICT Adoption by Indian Women: An analysis of two cases 346
Anindita Paul and Deepa RAY

Survey of fall detection techniques based on computer vision 351
Nuth Otanasap and Poonpong BOONBRAHM

SVM Approach for Human Gender Recognition from Video Sequences 356
Ken Nishio and Masao IZUMI

Sentiment Polarity Estimation Using Term Chunking and Numeric Phrases in Fact Sentence 359
Hidekazu YANAGIMOTO

Text Classification Tool for Sentiment Analysis on Social Media Monitoring in Retail Business 363
Vasutan Tunbunheng and Datchakorn TANCHAROEN

A Novel Interface of Interactive Evolutionary Computation for Automatic Sentence Generator 367
Kiyohito FUKUDA, Naoki Mori and Keinosuke MATSUMOTO

Robust Scheduling by Hybrid Method of Policy Tree and Estimation of Distribution Algorithm and Application to Resource Constrained Project Scheduling Problem with Rework Uncertainty 371
Jing TIAN, Xinchang Hao and Tomohiro MURATA

Human Detection from Surveillance Cameras with Considering the Influence of Shadows 378
Masao IZUMI

Automatic Update Method of WiFi Indoor Positioning Model using Pedestrian Dead-Reckoning Technique 382
Taniuchi Daisuke and Maekawa TAKUYA

Indoor Positioning System using Ultrasonic Sensors as a Position Information Infrastructure for a Wide Area 386
Chokatsu YARA, Masaki AKIYAMA, Shigenori Ioroi and Hiroshi TANAKA

Identifying Physical Traits with Body-worn Accelerometers: A Preliminary Investigation 390
Akira Masuda and Takuya MAEKAWA

From Linked Data to Linked Services: Developing Information Services for Sustainable Buildings 394
Vlado Stankovski and BoÅ¡tjan BRANK

The Development of Web Questionnaire with Business Process Modeling for Service Evaluation 400
Hisashi Masuda and Wilfrid UTZ

A study on a generic development process for SOA design and implementation 404
Toshimi MUNEHIRA

Continuous Improvement Approach of Software Process Based on Product Quality Inspection 412
Takahiro SEINO, Motoi YAMANE, Koichi Hayashi and Noriaki IZUMI

User-Customizable Database Access Application on Smart Devices for Small Businesses 416
Daisuke KIKUCHI, Toshiaki Ito and Masaru TEZUKA

Toward an Interactive Query Expansion: Concept driven and user feedback model 422
Krich NASINGKUN, Boontawee SUNTISRIVARAPORN, Thepchai Supnithi and Mitsuru IKEDA

Color-based Ranking Method for Web documents 430
Yurika MORIWAKI, Hiroshi Honda and Kosuke TAKANO

Assessing E-Business Deployment Framework for Thai SMEs: - The Case of Thai Restaurants in Melbourne, Australia - 434
Suttisak Jantavongso and Kulthon KASEMSAN

The actual situation about Practical Capabilities of Information Control in the Certified Nurse Administrator education first and second level trainees 441
Takako Izumi and Yukie MAJIMA

A Benchmarking Study of Standard Frameworks for Information Technology Governance 447
Wilai Sunthonwutinun and Vasin CHOOPRAYOON

Thai ICT HR Competency Standard Conceptual Framework for Supporting Sustainable Smart Thailand Towards AEC 2015 454
Kriengkrai BHUVANIJ, M.L.Kulthon Kasemsan and Prasong PRANEETPOLGRANG

A Comparison of the Need and Readiness of Laptops and Tablets Usage for Working at Tesco Lotus and BAAC Bank in Thailand 462
Sanon Chimmanee and Sunya CHIMMANEE

A Case Study: Thailand Customs Electronic System (TCES) in Aranyapraathet Customs House 468
Savika Nak-on and Monruthai RADEEROM

The Study of Need and Factors Affecting the Selection of Tablet between IPAD and Samsung Galaxy Tab compared between the Flight of Army Aviation Center's Pilots and Rangsit University's students 473
Pattama NOUMTHANONG, Sanon Chimmanee and Sunya CHIMMANEE

A preliminary proposed curriculum support system for information technology specialists based on the service science approach 479
Wooseok DONG, Kunio Shirahada and Michitaka KOSAKA

How do Sharing Service Providers Create Value? 488
Youji Kohda and Kazuhiro MASUDA

Service Value Co-creation for Enterprise IT Solution Services 493
Yukiko Nishioka and Michitaka KOSAKA

A Context-aware Information in Smart Home for Health Recommendation Service based on CARE Architecture 501
Konlakorn WONGPATIKASEREE, Azman Osman Lim and Yasuo TAN

A Proposal of KOSA Model for Service Innovation 509
Michitaka KOSAKA, Hiroyuki SAKANO, H.M. BELAL, Qi Zhang and Kunio SHIRAHADA

Value creation, co-production and co-creation in NGO-operated clothing industry 515
Md. Abul Kalam SIDDIKE, Amna Javed and Youji KOHDA

Capacity Maximization via Selection of Uniformly Distributed User for Uplink MIMO Communication over Î·-Î¼ Fading Channel 521
Theerasak SANGYAM, Chalie CHAROENLARPNOPPARUT, Prapun SUKSOMPONG, Kamol Kaemarungsi and Kazuhiko FUKAWA

An Analysis of Major Influencial Factors for Innovation in Thai Automotive Industry 526
Narongsak PONGSATHORNWIWAT, Chawalit JEENANUNTA, Thanaruk Theeramunkong and Hyunh VAN-NAM

Developing an Appropriate Data Collection Process on Traffic Accident in Thailand 536
Nitipon TANSAKUL, Suthathip Suanmali and Veeris AMMARAPALA

Innovation of Automobile Dealers in Hokkaido 544
Tomomi KANEKO, Wataru HASE, Setsuko Takahashi and Ryoju HAMADA

An approach of Software Development Support in Reference to Organization Management Paradigm 549
Noriko Wada and Michitaka KOSAKA

Building Knowledge of Customer Distinction for Safe Transaction in the Mail Order Industry 557
Hiroaki AZUMA, Masakazu TAKAHASHI, Kenichi Yoshida and Kazuhiko TSUDA

Active Contours (Snakes) and their Applications to Medical Images. A Mini-Lecture 564
Stanislav S. MAKHANOV

Optimization of 5-axis machining with application to dental implants 575
Samart Moodleah and Stanislav S. MAKHANOV

A Study of Retinal Blood Vessel Clustering for Finding the Main Vessel Convergence 581
Pongsate TANGSENG, Nittaya MUANGNAK, Pakinee AIMMANEE, Stanislav S. Makhanov and Bunyarit UYYANONVARA

Comparative Study of Initial Seed Point Required Approaches with the use of Vessel Distance Hint for OD Boundary Detection 587
Nabwan PRUKPAIBOON, Sarunkorn CHOTVIJIT, Nittaya Muangnak and Pakinee AIMMANEE

Top-view Based People Countingâ€¨Using Mixture of Depth and Color Information 593
Chonthisa Wateosot and Nikom SUVONVORN

An Image Segmentation Method for Glaucoma Detection using the CDR 597
Wanicha RUENGKITPINYO, Toshiaki Kondo and Waree KONGPRAWECHNON

Automatic Initialization of Multiple Active Contours Applied to Ultrasound Images of Breast Cancer 601
Khwunta KIRIMASTHONG, Stanislav S. Makhanov and Annupan RODTOOK

Multiple Quadratic Snakes for Segmenting Tumors in Medical Images 605
Yoichi Nakaguro and Stanislav MAKHANOV

Product Recommendation Method based on Onomatopoeia Expressing Texture 609
Ryuichi DOIZAKI, Saki IIBA, Takashi ABE, Takayuki Okatani and Maki SAKAMOTO

Method of evaluating metal textures by the sound symbolism of onomatopoeia 618
Junya YOSHINO, Akira YAKATA, Yuichiro SHIMIZU, Masaharu Haginoya and Maki SAKAMOTO

A Basic Study of Hot-and-cold Blower for Realizing Onomatopoeias 625
Yasushi Hamamura and Mitsunori MATSUSHITA

Why are specific onomatopoeias evoked from specific images? - Investigating correlations between image features and quantified onomatopoeias 631
Takanori Komatsu and Shoma INAOKA

Constructing a Knowledge Base of Onomatopoeia-Action Relations for an Interactive Illustrated Storybook 635
Hiroki UEMA, Daiki Hashui and Mitsunori MATSUSHITA

Comparison of onomatopoeia use among areas using Diet records of Japan 641
Sachiko HIRATA-MOGI, Satoshi NAKAMURA, Takanori Komatsu and Kimi AKITA

Risk-aware Data Replication to Massively Multi-sites against Widespread Disasters

Shinya Matsumoto† **Takaki Nakamura‡** **Hiroaki Muraoka‡**

†Yokohama Research Laboratory, Hitachi, Ltd.
shinya.matsumoto.gn@hitachi.com
‡Research Institute of Electrical Communication, Tohoku University
{takaki,muraoka}@riec.tohoku.ac.jp

Abstract

Immediately after a widespread disaster, critical information such as census registers should be available in the disaster-affected area to confirm people's safety. However, when a widespread disaster actually happened, such information was not available in the Great East Japan Earthquake of 2011 even though the information were protected in safe and remote site over the Internet. This is because the disaster damages both data and internet access of the area. To enable the information available even immediately after a widespread disaster occurs, we propose Risk-aware Data Replication (RDR), which replicates data to safe and nearby sites in the area without the wide area network. It evaluates the site-to-site risk of data loss, and seeks low-risk sites from massively multiple sites using mathematical programming approach. The result of a simple earthquake simulation shows that RDR improves data availability to more than 10% compared to the case of random selection of replication sites and the computation time is less than several seconds even under the 1350 sites. Our contribution is to provide an organized, effective, and feasible way to decide replication sites based on their disaster risks under massively multiple sites environment.

Keywords: Storage System, Availability, Replication, Backup System, Disaster Recovery

1 Introduction

Many kinds of natural disasters intermittently occur all over the world. A widespread disaster, such as an earthquake and tsunami, may damage information systems seriously and stop their services. However, immediately after a widespread disaster, local governments and medical institutions need critical information: census registers for confirmation of safety, medical records, and prescriptions for health-care. Therefore, our research purpose is to continue such services in a disaster-affected area.

Remote replication technique, which replicates data into a remote site over the Internet, is widely used to mitigate such disaster effects on a service. For example, SnapMirror [1] transfers data of a primary site to a remote replication site which is never damaged simultaneously, and the replication site provides the primary's services instead when a disaster damages them.

Remote replication technique is not sufficient when a widespread disaster damages both a primary site and internet access of the disaster-affected area. This is because the recipients of services in the area cannot access the replication site over the Internet. This problem actually happened in the Great East Japan Earthquake of 2011.

Therefore, our goal is to design a highly-available storage system that enables recipients of services in disaster-affected area to access the alternative services immediately after a widespread disaster. To achieve this, we propose Risk-aware Data Replication which utilizes geographically-distributed sites which are usual in branch offices of local governments and medical institutions, and evaluates the site-to-site risk of data loss, and seeks low-risk sites using mathematical programming approach.

2 Related Work

Effective data distribution for availability has been widely studied in storage systems communities.

Dynamo [2] distributes data at a set of nodes (i.e., storage hosts) based on consistent hashing technique [3]. Dynamo system determines rep-

lication targets by using a "ring" created by the output range of circular hash function. Each node is assigned to the random value within the ring space, and replicates data at its successors. Dynamo distributes data essentially in a random manner because it partitions hash range randomly in adding a new node to the system, and never considers the node's safety against widespread disasters. Therefore, Dynamo requires high redundancy of data if we hope for high availability against such disasters.

Cassandra [4] also distributes data at a set of nodes based on consistent hashing technique. Additionally, it provides various replication policies such as "Rack Aware" and "Datacenter Aware". By activating these policies, Cassandra system replicates data at different racks or different datacenters from the node storing primary data. These policies enable the system to avoid failures due to power outage, cooling failures, network failures, and natural disasters. Cassandra only considers whether the replication target is installed in the same rack or the same datacenter as the primary node, and never considers the node's safety against widespread disasters. Therefore, Cassandra also requires high redundancy of data if we hope for high availability against such disasters.

3 Problem Establishment

3.1 Target System

Our target system consists of branch offices of local governments and medical institutions, which can be considered as a class of geographically-distributed storage systems illustrated in Figure 1.

It consists of some areas. Each area is distant enough from each other not to be damaged simultaneously by an earthquake, and is connected with a Wide Area Network (WAN).

Every area has some sites connected with an Local Area Network (LAN) in it. Every site includes servers and a storage system, and provides its own services to recipients in the area. The number of sites may be more than several thousand because there are such numbers of medical institutions in a large prefecture of Japan [8].

We assume that an earthquake damages only a part of an area, that is, some of sites, a part of LAN, and a WAN connection, but each sites in the same area can communicate with each other

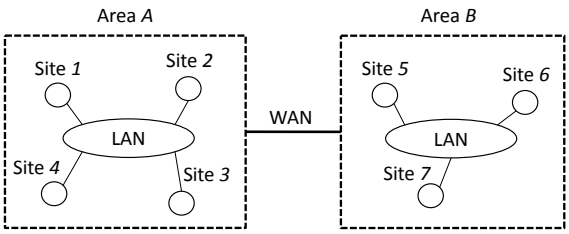


Figure 1. Target System

using the area's LAN.

The target system improves data availability using the following way: each site replicates its data to other sites periodically; when an earthquake damages a part of an area; one of the surviving site takes over the services of the damaged site.

3.2 Requirements

To design the highly-available storage system under the target system, we should meet two requirements: a protection requirement and an accessibility requirement.

The protection requirement is that a primary site and its replication sites should not be simultaneously damaged by an earthquake. This requirement is obviously necessary not to lose both primary data and its replica. Conventional remote replication technique meets this requirement by replicating data to a site in other area over the Internet.

The accessibility requirement is that a primary site and its replication sites should be in the same area. This is because recipients of services in disaster-affected area cannot access its replication site in another area when an earthquake damages WAN connection.

3.3 The Issue

The issue we have to solve is to meet both the requirements above: to find pairs of two sites that are not damaged simultaneously in the same area. To meet both the requirements, there are at least two difficulties.

The first difficulty is that it is precisely unpredictable where the next earthquake occurs and how strong it is. It is rather easy to find a safe site against a particular earthquake whose location and strength are known because we can just replicate data to a sufficiently far site from the earthquake. However, the unpredictable earthquakes may cause data loss by damaging both primary data and its replicas simultaneously.

The second difficulty is that the number of

combinations with pairs of primary-replication sites may be huge. In conventional remote replications, the number of replication sites is no more than three sites, therefore, an administrator can select a replication site easily. By contrast, in our research, the number of replication sites may be more than several thousands, and the number of combination is much more, therefore it is difficult for an administrator to check all the combinations and select a proper replication site.

4 Our Approach

Replicating data to all the sites is one of the simple solutions. However, it requires a lot of storage capacities. If the storage system consists of a thousand sites and if each site needs a storage capacity of 1TB for its services, this simple solution requires each site to have its storage capacity of 1PB.

Therefore, we propose Risk-aware Data Replication (RDR). RDR quantifies and evaluates an earthquake risk of every site, and only replicates data to low-risk sites. Therefore, it can save their storage capacities with maintaining the data availability.

We construct two design policies of RDR to overcome the two difficulties described in Section 3.3.

The first policy is that RDR focus on multiple and stochastic earthquakes information such as earthquake forecast maps provided by a government agency. This policy overcomes the first difficulty, and makes RDR avoid unpredictable data loss.

The second policy is that we use the model known as mathematical optimization problem. This policy makes RDR capable to support massively multiple sites because the model provides organized procedure to decide pairs of primary-replication sites.

5 Risk-aware Data Replication

To put RDR with the two policies described in Section 4 into action, we design it with following three steps: modeling an earthquake risk, formulating the problem into the form of mathematical optimization problem, and solving it numerically to decide safe pairs.

5.1 Modeling an Earthquake Risk

In this paper, we introduce a data loss probability

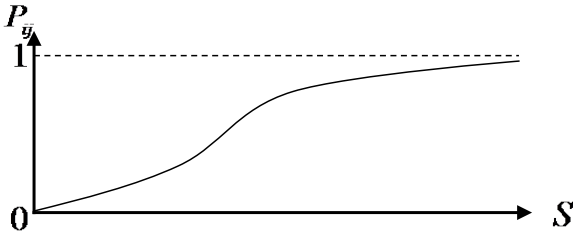


Figure 2. The Property of P_{ij} to Model

which indicates two different sites are simultaneously damaged in the same area due to target earthquakes in order to model the earthquake risk. This is possible since the target earthquakes are stochastic in the earthquake forecast map and the risk can be described by a form of probability.

Here, we model the data loss probability under an assumption that it is equally probable where earthquakes occur and how strong it is for the simplest case of the earthquake forecast map.

Let S denote the strength of an earthquake, the data loss probability P_{ij} between sites i and sites j satisfies the following two properties described in Figure 2:

A) P_{ij} is continuous and increases as S increases,

B) P_{ij} is positive and converges to 1 as $S \to \infty$,

where $i, j \in \mathbb{N}$, $i, j \in [1, n]$, $\mathbb{N}$ is a set of natural numbers, and $n \in \mathbb{N}$ denotes the total number of sites. These properties follow our intuitions. The property A indicates that the two sites are more likely to be damaged as the strength increases. The property B describes that all the value of the probability is more than 0 and less than 1.

One of the descriptions of the data loss probability with these intuitive properties can be derived in the following steps:

1. calculates a strength of an earthquake S from the given value d_{ij}, which denotes the distance between sites i and sites j, using the relation of

$$S(d_{ij}) = -\log_{10} d_{ij};$$

2. calculates P_{ij} by using the sigmoid function ζ with given parameters a, b of

$$P_{ij}(d_{ij}) = \zeta\left(a\left(-S(d_{ij}) - b\right)\right) = \frac{1}{1 + e^{a(\log_{10} d_{ij} - b)}}.$$

The first step is to describe the reasonable strength S, and is based on an statistical earth-

quake model [6]. We see the maximum speed of ground as the strength of earthquake, and assume that the increasing rate of the probability depends on the decreasing rate of the speed. The second step is to satisfy the properties A and B. Using the sigmoid function ζ with $a > 0$, it assures $P_{ij} > 0$ and P_{ij} converges to 1. Additionally, we use two design parameters a, b to tune the curve of P_{ij}. P_{ij} converges to the step function as $a \longrightarrow \infty$, and we can model P_{ij} with a gradual slope if we set a small a. Essentially, these parameters depend on the property of target earthquakes, however, the design method remains future work.

The derived data loss probability follows general intuition that a strong earthquake is likely to damage two sites which are located far from each other, and that the two sites near each other are likely to be damaged simultaneously.

5.2 Formulating into the Mathematical Optimization Problem

We formulate the problem to decide replication sites into the Integer Programming Problem (IPP) which is one of the mathematical optimization problems. The IPP consists of an objective function and constraints including integer variables.

We aim to get a combination with pairs of the primary-replication sites that optimize the data amounts surviving from an earthquake. By using P_{ij} as a weight coefficient of an objective function to decide replication sites, we encourage RDR to decide safe sites as a replication sites.

We also consider the number of replicas to create and the amount of storage capacity that can be consumed by replicas. If the number of replicas is large, the availability after an earthquake is clearly large, but the storage capacity consumed by replicas is also large. To make this trace-off tunable, the number of replicas should be a design parameter. The amount of storage capacity for replicas is valuable and limited resources for many systems. If the amount is large, the replicas would concentrate in the safest site. This is not necessarily beneficial. To avoid such a case, the storage capacity for replicas should be a design parameter. These two conditions can be described into constraints of the IPP.

We show the formulation of the IPP below.

(1) Objective Function

RDR aims to protect data of sites in an area against target earthquakes. Here, we aim to minimize the total amount of data which is damaged and lost. Thus, we set an objective function of

$$f(x_{12}, \dots, x_{n(n-1)}) = \sum_{i=1, i \neq j}^{n} \sum_{j=1}^{n} D_i P_{ij} x_{ij},$$

where $x_{ij} \in \{0, 1\}$ denotes whether a site j has a replica of a site i or not, and $x_{ii} = 0,\ \forall i$. D_i denotes the primary data amounts in site i, and P_{ij} denotes the data loss probability described in Section 5.1. This definition of f denotes the total amounts of data which are expected to be damaged for a combination of the variables x_{ij}. The combination of x_{ij} to minimize f is the most highly available solution against target earthquakes.

(2) Constraints

RDR uses two constraints: a redundancy constraint and a storage capacity constraint.

The redundancy constraint is to regulate the number of replicas to create from a primary site in a replication process. This constraint is necessary because every site tries to create as many replicas as possible without it. Thus, the redundancy constraint is set to bound the maximum value of data redundancy, and described by

$$\sum_{j=1}^{n} x_{ij} = R_i, \forall i,$$

where R_i denotes the number of replicas of site i, and is given by an administrator.

The storage capacity constraint is to regulate the number of replicas that a replication site receives in a replication process. This constraint is necessary because every site has finite storage capacities. Thus, the storage capacity constraint is set to bound the maximum value of storage capacity consumed by replicas, and described by

$$\sum_{i=1}^{n} D_i x_{ij} \leq F_j, \forall j,$$

where F_j denotes the number of storage capacity of site j and is given by an administrator or each sites.

5.3 Solving the IPP

To minimize f, we can use an existing optimal algorithm. However, IPP generally takes time to solve. Thus, we use the following two algo-

rithms.

(1) Branch-and-Bound method (BB method) [7]
BB method is an algorithm used to seek the combination of x_{ij} which minimize the objective function f. This method guarantees that the solution is optimal, but the computation time generally increases fast as the total number of sites increases.

(2) Greedy method
Greedy method is an algorithm that is used to search for the combination of x_{ij} by finding the component of x_{ij} which has the largest P_{ij} in turns. This method does not guarantee an optimal solution, however, the computation time is generally small.

6 Implementing RDR Simulator

We implemented RDR simulator by C to evaluate RDR. Figure 3 describes an input-output block diagram of it. The RDR simulator consists of three parts of computation: a field creator, a risk creator, and a pair creator.

The field creator simulates sites in an area, which has a flat ground, a slope, or a mountain, and outputs a field information and sites information. The field information includes the geological formation of the ground described by specified coordinates of computational mesh. The sites information includes their locations, their usage and free amounts of storage capacities.

The risk creator calculates the site-to-site risk P_{ij}. The risk creator receives two pairs of (d_{ij}, P_{ij}) as a risk hint, and calculates the parameters a, b with linear interpolation in the input variable space of the sigmoid function ζ. Then it outputs the risk information, the value of probabilities of all the pairs of sites.

A pair creator decides a combination with safe pairs of primary-replication sites. It formulates and solves the IPP by using constraints information and a specified algorithm. The constraints information includes data redundancy. The algorithm can be selected from BB method and Greedy method. For BB method, it invokes lp_solve 5.5, and for Greedy method, we just implemented its algorithms. Finally, it outputs the safe pairs.

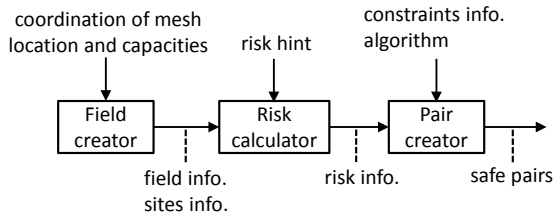


Figure 3. RDR simulator

7 Evaluation

We measured the RDR's improvement of availability and computation time to decide pairs using a simple earthquake simulator, to evaluate RDR's effectiveness and feasibility.

7.1 Earthquake Simulator

We implemented a simple earthquake simulator to evaluate RDR's improvement on availability. The earthquake simulator simulates to damage sites in a field.

We input the earthquake's magnitude and the location of its hypocenter including the depth to the simulator. The simulator calculates the damage probability of every sites based on the statistics of an earthquake [6][8], using the probability, the simulator decides if sites are damaged or not.

7.2 Simulation Setup

We conducted the simulation on a server with a processor (2-core 1.87 GHz, Xeon E5502; Intel Corp.) and 6 GB RAM, and set parameters of the RDR simulation and the earthquake simulation in the following.

The field is flat and 200km $\times$ 200km wide. The sites are located randomly on the field, and the number n is 135 or 1350. Each site has only one data and only one storage capacity for replica, that is, $D_i = 1, F_i = 1, \forall i$, and the replicates its data only one time, that is, the redundancy $R_i = 1, \forall i$.

We simulated two kinds of earthquake to evaluate robustness of the algorithm. One is a far-field earthquake, the hypocenter of which is far away from every site. To simulate this, we set an earthquake of which hypocenter is $(400, 100, -40)$ and its magnitude is 8. Another is a near-field earthquake, the hypocenter of which is near some sites. To simulate this, we set an earthquake of which hypocenter is $(100, 100, -40)$ and its magnitude is 8.

We set (10, 0.9) and (50, 0.5) for the risk hint described in Section 6, and repeat each kind of earthquake 500 times to obtain an average of measured data.

We use not only the BB method and the Greedy method described in Section 6.3, but also Random method for algorithms in order to compare. The Random method decides pairs of primary-replication sites in a random manner, and the expectation of availability is ideally 75% in the case that the number of replicas is 1 and damaged half of the total number of sites.

7.3 Results

(1) Distribution of Primary-replication Pairs
Figure 4 shows the histogram of distances between two sites each of which has the other's replica when the number of sites is 135.

In the case of the BB method, the distribution is concentrated to the distance around 160, there are no pairs except 140-180, and the average of the distance is 154.6. This shows that BB method pairs sites to biasing toward its average of the distribution.

It shows that Greedy method pairs sites which are more distant than Random method and BB method do. There are pairs of which distance is over 180 and near the maximum value of the distance 259.2. This shows that the Greedy method pairs sites which have as long distance as possible.

(2) Availability
Figure 5 shows the available data ratios, the percentage of data amounts that survive from an earthquake when the total number of sites is 135. The values of the BB method is more than 20%, larger than that of the Random method, but the values of the Greedy method is less than 20% smaller in the case of near-field earthquake. This shows that RDR is effective to improve data availability if the algorithm is well-designed. For a far-field earthquake, the Greedy method value is 0.6% lower than that of BB method. This shows the Greedy method to be as effective as the BB method under a far-field earthquake.

(3) Time to Decide Safe Pairs
Figure 6 shows the time to decide pairs. This shows that the Greedy method can select replication sites in a shorter time than the BB method does, and that the time is less than several seconds.

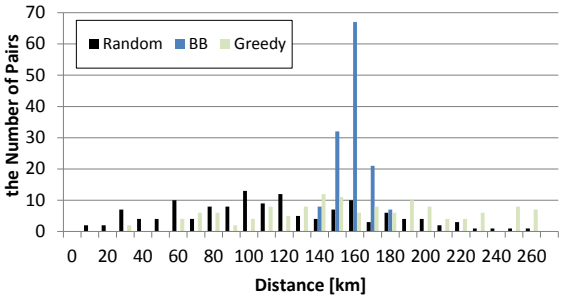


Figure 4. Histogram of Site Distance

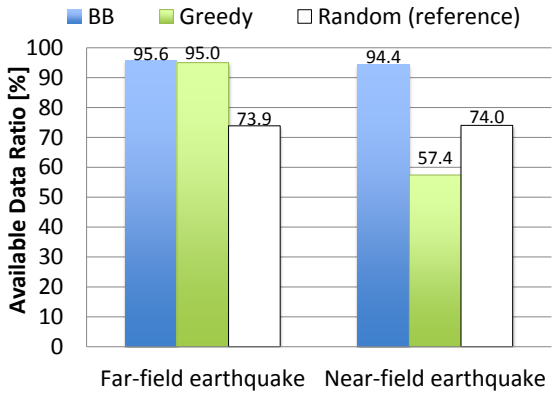


Figure 5. Availability

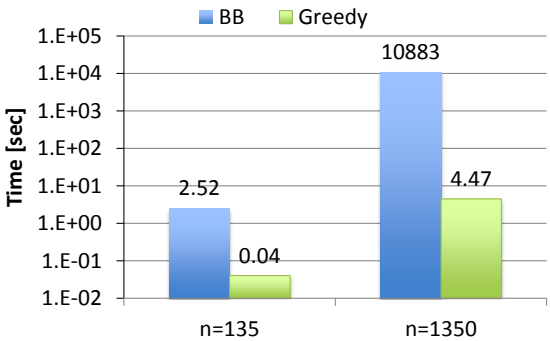


Figure 6. Time to Decide Pairs

8 Conclusions

This paper presents Risk-aware Data Replication, which replicates data to nearby and safe sites based on evaluating earthquake risks. We show that a problem to decide replication sites can be formulated into a form of an Integer Programming Problem model by using a probability based on a site-to-site risk against an earthquake, and that we can solve it using existing algorithms. By simulating an earthquake, we show that, depending on the algorithm, RDR improves data availability to more than 20% larger than the case of random selection of replication sites, and the

computation time is less than several seconds even under the 1350 sites.

In the future, we plan to improve the definition and design method of a site-to-site earthquake risk. The proposed site-to-site risk seeks a weak solution against the earthquakes that occur between the primary and replication sites. To improve this, we plan to derive the site-to-site earthquake risk from an earthquake forecast map which predicts a geographical range of a hypocenter mathematically.

Acknowledgment

We thank Masatoshi Shimbori, Hiroshi Ichinomiya, and Mitsuyoshi Igai for research computing support, and Mitsuo Hayasaka, Atsushi Sutoh, and all the members of our project team for their useful comments. This work is supported as “Research and Development on Highly-functional and Highly-available Information Storage Technology”, sponsored by the Ministry of Education, Culture, Sports, Science and Technology in Japan.

References

[1] NetApp, Storage Infrastructure for Cloud Computing, *http://www.netapp.com/us/media/wp-7081.pdf*, 2009.

[2]G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. Vogels, Dynamo: Amazon's Highly Available Key-value Store, In *Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, pp.205–220,* 2007.

[3]Karger, D., Lehman, E., Leighton, T., Panigrahy, R., Levine, M., and Lewin, D., Consistent hashing and random trees: distributed caching protocols for relieving hot spots on the World Wide Web, In *Proceedings of the Twenty-Ninth Annual ACM Symposium on theory of Computing (STOC '97). ACM Press, New York, pp. 654-663*, 1997.

[4] A. Lakshman, P. Malik, Cassandra: A Decentralized Structured Storage System, *ACM SIGOPS Operating Systems Review*, vol. 44, no. 2, pp. 35-40, 2010.

[5] Statistics Japan, Medical care institutions and pharmacies by prefecture, *http://www.stat.go.jp/data/nenkan/zuhyou/y2120000.xls*, 2010.

[6] H. Si, S. Midorikawa. Attenuation Relationships of Peak Ground Acceleration and Velocity Considering Effects of Fault Type and Site Condition, In *Proceedings of the 12th World Conference on Earthquake Engineering (WCEE 2000), CD-ROM, No.532,* 2000.

[7] E. L. Lawler and D. E. Wood, Branch-and-Bound Methods: A Survey, *Operations Research July/August, vol.14, no.4, pp.699-719,* 1966.

[8] J. Miyakoshi, Y. Hayashi, K. Tamura, N. Fukuwa, Damage ratio functions of buildings using damage data of the 1995 Hyogo-Ken Nanbu earthquake, In *Proceedings of the 7th International Conference on Structural Safety and Reliability (ICOSSAR ‘97), vol.1, pp.349-354*, 1997.

Evaluation of SSD Tier Method and SSD Cache Method in Tiered Storage System

Shinichi Hayashi† **Norihisa Komoda‡**

†Hitachi, Ltd., Japan
‡Osaka University, Japan
shinichi.hayashi.ez@hitachi.com
komoda@ist.osaka-u.ac.jp

Abstract

We evaluated storage I/O response time with the SSD tier method and the SSD cache method on a tiered storage having SSD and HDD. The simulation result shows the SSD tier method improves the storage performance in write-intensive and low I/O locality workload with small SSD capacity. The SSD cache method is effective in read-intensive and high I/O locality workload with large SSD capacity. Utilizing both methods simultaneously is effective with write-intensive workload. The storage I/O response time is reduced by up to 28% utilizing both methods simultaneously compared to using the SSD tier method or the SSD cache method separately.

Keywords: Tiered Storage, Tier, Cache, I/O Response Time, SSD

1 Introduction

With recent improvements in information technology, the amount of data retained by companies has increased exponentially. The capacity of HDDs (Hard Disk Drives) has continued to increase; however, the performance of these devices has not improved significantly. Therefore, HDDs can potentially become bottlenecks. SSDs (Solid State Drives), which are much faster than HDDs, are currently attracting attention. When HDDs are bottlenecks, replacing HDDs with SSDs could increase performance. However, SSDs generally are more expensive and have less storage capacity, therefore, budget and capacity constraints limit a company's ability to replace all HDDs with SSDs.

In General, I/O (Input and Output) activities have a locality according to purposes and characteristics of data. Therefore, if frequently accessed areas are stored on a SSD and rarely accessed areas are stored on a HDD, I/O performance will increase and system cost will be reduced. A frequently accessed area is denoted as a hot area and a rarely accessed area is denoted as a cold area.

The followings are the 2 methods to utilize SSD on the storage having both SSD and HDD.

(1) Method dividing and locating data to SSD or HDD

(2) Method locating all data to HDD and coping partial data to SSD

In this paper, we call the first method, SSD tier method and the second method, SSD cache method. The SSD tier method and the SSD cache method are evaluated respectively in several researches [1][2][3]. Several companies provide the storage which has SSD tier method or SSD cache method [4][5][6][7][8].

The SSD tier method improves I/O performance by allocating hot areas to SSD. However, because this method optimizes data location every 1 hour to 24 hours, when the hot areas move, it takes 1 hour to 24 hours to improve the I/O performance. Since the SSD cache method caches accessed area to SSD immediately, it improves the second and subsequence accesses to the area. When cached data is ejected frequently, however, the SSD cache method needs to write the data to HDD, therefore, there is a possibility to increase HDD load which leads to lower storage I/O performance.

We consider that utilizing both methods simultaneously addresses the problems above. We assume that appropriate number of SSDs for tier and appropriate number of SSDs for cache may differ on the basis of the number of I/O, read-write ratio, and I/O locality. In this paper, we evaluate the effects of these methods by I/O simulation and clarify conditions when these

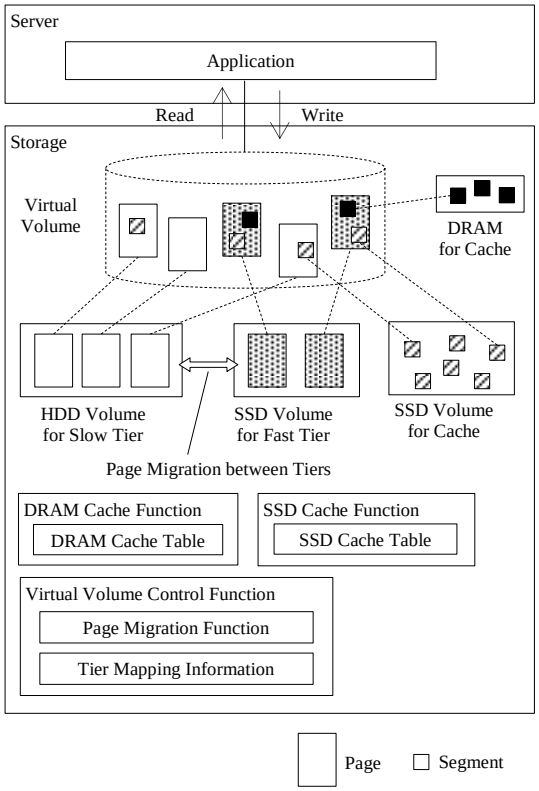


Figure 1. Overview of Storage System

methods are effective.

2 SSD Tier Method and SSD Cache Method

2.1 Overview of Storage System

Figure 1 shows the overview of the storage system which we use in this paper. The storage consists of virtual volumes which servers access, a virtual volume control function which controls the virtual volumes, a DRAM (Dynamic Random Access Memory) cache function, a DRAM, SSD volumes, and HDD volumes. The virtual volumes consist of areas called pages. SSD volumes as a fast tier or HDD volumes as a slow tier are assigned to each page.

The storage system utilizing several tiers which have different performance is referred to as a tiered storage. The virtual volume control function manages the mapping information and reads from fast or slow tier, or writes to fast tier or slow tier on the basis of access from applications on servers to the virtual volumes. All data on the virtual volumes should be thus allocated to fast or slow tier.

2.2 SSD Tier Method

The SSD tier method is designed to improve I/O performance by allocating SSD to frequently accessed page. The tier control function measures the number of I/Os (n_i) in page i for a certain period, which is defined as an I/O measurement period (p [hours]), ranks the pages in order of the number of I/Os ($1 \leq r_i \leq N$), and determines $NewTier_i$ of page i on the basis of the formula (1). In the formula (1), 1 means the fast tier, 2 means the slow tier, N means the number of pages in the storage, and h means the rate of the fast tier capacity to the slow tier capacity (fast tier capacity rate).

$$NewTier_i = \begin{cases} 1 & (r_i \leq Nh) \\ 2 & (r_i > Nh) \end{cases} \quad (1)$$

Following the tier determination, when $NewTier_i$ is different from the current tier ($Tier_i$), a page migration function migrates the data from $Tier_i$ to $NewTier_i$. We call this process data migration between tiers. Since the fast tier is more expensive, all areas of the fast tier are allocated to pages. The virtual volume control function measures I/Os then determine the page tier and the page migration function controls the migration between tiers.

2.3 SSD Cache Method

The SSD cache function stores accessed data to SSD which has high performance media to accelerate the subsequent I/Os. The SSD cache function caches the data with a smaller unit called segment.

When the storage receives a read request from the server, the SSD cache function refers a SSD cache table to check requested data is in the SSD. If the data exists in the SSD, it reads the data from the SSD and sends a result to the server. Otherwise, it sends the read request to the fast or slow tier, sends a result to the server, and stores the data to the SSD. At this time, if there is no free space on the SSD to cache the data, it ejects some data. Several algorithms such as Least Recently Used (LRU) are proposed to determine which data to be ejected from the SSD. We define SSD cache hit as a requested data existing in the SSD when the SSD cache function receives a request from the server. We also define SSD cache miss as a requested data not existing on the SSD.

When the storage receives a write request from the server, the SSD cache function stores the data to a free space on the SSD and sends the completion of the write command to the server. If there is no free space on the SSD, the SSD cache function ejects some data from the SSD before the process above. The SSD cache function records a state that the data is not yet written in the fast or the slow tier. When the SSD cache function ejects from the SSD, it writes the cached data in the fast or the slow tier. Since storage system has a redundancy for data protection, we discuss a write-back cache in this research.

The SSD cache function resides between the DRAM cache function and the fast and slow tier. The DRAM cache function receives I/Os from the server and stores data which it accesses once to DRAM for cache. The DRAM cache function is the same as the SSD cache function except caching data to DRAM.

DRAM has smaller capacity with higher price, on the other hand, SSD has larger capacity with low price. Therefore, the capacity of the SSD cache is larger than the DRAM cache. Since SSD is a non-volatile device, it is possible to utilize SSD as the higher tier as well as cache function.

3 Comparison of SSD Tier Method with SSD Cache Method

3.1 Comparison of Characteristics

Table 1 shows characteristics of the SSD tier method and the SSD cache method. The SSD tier method migrates data which needs to be arranged in SSD from HDD. On the other hand, The SSD cache method copies data from HDD to SSD. Since the SSD cache method has duplicated data, the SSD tier method provides a better efficiency. When the number of I/Os to pages decreased, the pages have to be migrated from SSD to HDD in the SSD tier method, whereas only segments which data is not written in the fast or the slow tier only have to be written in the SSD cache method.

In the SSD tier method, data control size is from 1 MB to 1 GB. Control cycle is from 30 minutes to 1 day, which slowly migrates large data. On the other hand, in the SSD cache method, the data control size is from a few KBs to a few hundred of KBs. Cache miss triggers the control and it frequently copies small data.

Table 1. Characteristics of SSD Tier Method and SSD Cache Method

	SSD Tier Method	SSD Cache Method
Data Control	Move	Copy
Data Control Size	Large (1 MB to 1 GB)	Small (A few KBs to a few hundred of KBs)
Control Trigger	Every 30 minutes to 1 day	Every Cache Miss
Removal Algorithm from SSD	LFU, etc.	LRU, etc.
Management Table	Covering all pages	Covering only cached area

Regarding data location optimization, the SSD tier method determines data location on the basis of the number of I/Os to each page. The SSD cache method determines data location on the basis of LRU algorithm, for example.

Regarding data location management method, the SSD tier method manages relationships between all pages in the virtual volumes and areas in the fast or the slow tier. Therefore, if the capacity that stores the page mapping table is limited, small size pages cannot be used. The SSD cache method manages only cached areas, therefore, small size segments can be used.

3.2 Comparison of I/O Performance

In the SSD tier method, due to its lower frequency of data location optimization, it takes time for I/O performance to improve when accesses intensively concentrate on a specific area. On the other hand, the SSD cache method can improve I/O performance immediately from the second access in SSD cache hit.

Since the SSD tier method migrates larger size data for data location optimization, when intensively accessed area moves within a same page, it enables the improvement of I/O performance from the first access. Also due to its less frequency in data location optimization, when the intensively accessed area moves and gets back to the same page, the SSD tier method allows SSD access from the first time. Furthermore, the SSD cache method must write the cached data that is not written in the fast or slow tier frequently when there are frequent copies to SSD and ejects from SSD.

4 Evaluation

4.1 Evaluation Items

The SSD tier method and the SSD cache method both have different merits which depends on I/O characteristics such as the number of I/O, read and write ratio, and I/O locality. We evaluate the effects of these two methods by simulating I/Os using I/O trace log captured in a real production environment. These two methods can be applied simultaneously. We also evaluate these effects by adjusting the number of SSD used as the fast tier and as cache as a parameter.

Information system needs to provide performance regulated by SLA (Service Level Agreement) to information system users. SLA regulates that system response time should be, for example, within 300 milliseconds for 99.9% of its requests [9][10]. Information system provider sets SLO (Service Level Objective) to meet SLA. If the system response time is regulated by SLA, the provider regulates storage response time by SLO, thus, we consider the storage response time as storage I/O performance by simulating I/Os in this paper.

4.2 Simulator

We developed a simulator using Perl language to compute storage I/O response time. It simulates a DRAM cache process, a SSD cache process, and a data migration between the fast and the slow tier.

Figure 2 shows the simulation model. It simulates I/Os on the basis of the I/O trace log. The I/O trace log includes I/O time stamps, LBA, I/O size, and I/O type (read or write). SSD and HDD have their own queue, and their response times depend on the I/O type (see section 4.3).

On the basis of the I/O trace log, the DRAM cache function handles the DRAM cache process. The DRAM cache process is described in section 2.2. In case of DRAM cache hit, the DRAM cache function accesses the DRAM and the simulator records the response time. In case of DRAM cache miss, the I/O is sent to the SSD cache function.

The explanation of SSD cache function is given in section 2.2. In the case of SSD cache hit, the SSD cache function identifies the device which contains the data and enqueues the I/O and then the simulator records the response time. In

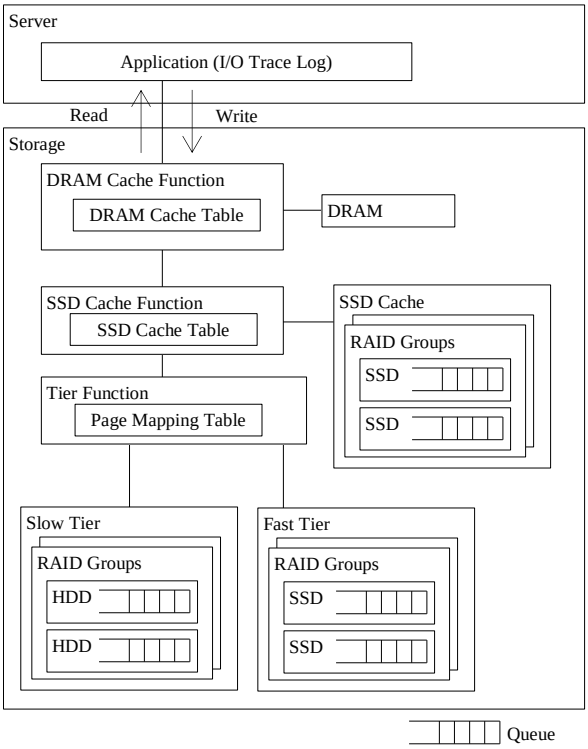


Figure 2. Simulation Model

the case of SSD cache miss, the I/O is sent to the tier function. The tier function then identifies the device that contains the data on the basis of the page mapping table and enqueues the I/O, and then the simulator records the response time.

The simulator also enqueues I/Os for ejecting the unwritten data from the DRAM or SSD cache to the fast or slow tier and I/Os for migration between tiers. The storage I/O performance does not include these I/Os since these are not from the application.

4.3 Simulation Condition

Table 2 shows an overview of the I/O trace logs [11] which are used in this simulation. These I/O trace logs (Financial 1 and Financial 2) are from online transaction processing (OLTP) applications running at two large financial institutions. Financial 1 contains a number of write requests and Financial 2 contains a number of read requests.

Figure 3 shows a distribution of the I/O trace logs. In Financial 1, for example, 57% of I/Os to total I/Os concentrate on 20% areas to total areas. In Financial 2, 84% of I/Os to total I/Os concentrate on 20% areas to total areas. Figure 3 indicates that Financial 2 has higher locality than

Table 2. Overview of I/O Trace Log

Name	Financial 1	Financial 2
IOPS	825	431
Read Rate	15.4%	78.5%
Average I/O Size	16.1 KB	18.8 KB
Volume Capacity	8.4 GB	8.4 GB
Record Period	12 hours 8 minutes	11 hours 23 minutes

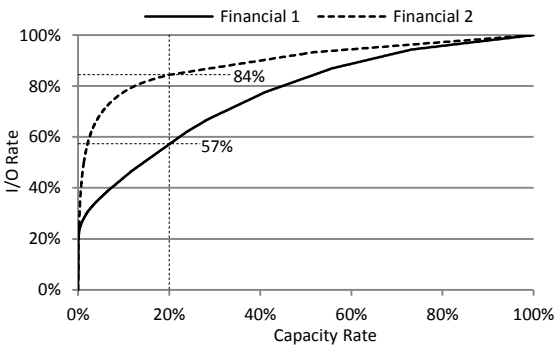


Figure 3. I/O Distribution of the I/O Trace Logs

Financial 1. In order to simulate a write-intensive with high locality workload and a read-intensive with low locality workload, we also simulate a condition that swaps read requests for write requests.

Figure 4 shows performance per device that is used in this simulation. Response time in each device to process a queued I/O is based on I/O type (read or write), I/O size, and device type. The values shown in Figure 4 are measured in the environment shown in Table 3. In order to prevent the effect of the DRAM cache function in the device performance measurement, the devices are mounted with the sync option. We used ext2 file system since ext3 and ext4 have a journaling overhead. DRAM read response time is 3.2 microseconds and write response time is 3.8 microseconds, which is measured in the same environment on Table 3.

Table 4 shows the number of devices and loads in the simulation. We define the rate in the number of SSDs to the number of HDDs as SSD rate. We adjust the SSD rate to 10%, 20%, and 30%. The number of SSDs and the number of HDDs are shown in Table 4. The maximum number of devices in this simulation is 1,500 on the basis of the maximum number of device [4][5][6][7][8]. In order to protect data, two devices are configured to a RAID 1. One I/O trace

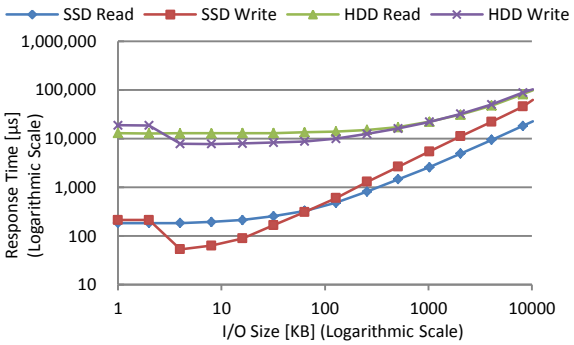


Figure 4. Device Response time

Table 3. Device Response Time Measurement Configuration

CPU	4 Cores, 2.13 GHz
Memory	12 GB
SSD	6 Gbps SATA, MLC, 240 GB
HDD	6 Gbps SATA, 7,200 rpm, 500 GB
OS	CentOS 6.4 (Kernel 2.6.32-358)
I/O Scheduler	noop
File System	ext2
Mount Option	sync
Benchmark Software	fio 2.0.14
I/O Type	Random Read and Write
Access Area	200 GB

log is simulated per one HDD RAID group. When the number of HDDs is 100, for example, 50 I/O trace logs are simulated to each RAID group simultaneously. We define SSD cache rate as the rate of the number of SSDs for the SSD cache method to the total number of SSDs. The other SSDs are controlled by the SSD tier method. In this simulation, we adjust the SSD cache rate from 20% to 80% in increments of 20%.

The page size is 10 MB in the SSD tier method MB and the size of segment is 8 KB in the SSD cache method. I/O measurement period is 1 hour. Because data migration between tiers affects I/O response time, the page migration function takes a data migration interval between tiers. It sets the transfer rate of migration between tiers to the HDD maximum transfer rate to 10% instead of migrating target page in sequence.

5 Result and Discussion

5.1 Comparison of SSD Tier Method and SSD Cache Method

Table 4. Number of Devices and Loads in the simulation

SSD Rate	# of SSDs	# of HDDs	Total
20%	250	1,250	1,500
30%	340	1,156	1,496

Figure 5 shows the simulation result of the storage I/O response time. The vertical axis shows the normalized storage I/O response time where the storage I/O response time with the SSD cache method is 100%. The horizontal axis shows SSD cache capacity rate to total capacity. F1 and F2 show the simulation condition in Financial 1 and Financial 2 respectively. "rw" means the condition that swaps read requests for write requests.

When we compare the SSD tier method only with the SSD cache method only, the storage performance gets better with the SSD tier method and F1-30% condition. From this result, it is considered that the SSD tier method is more effective with write-intensive, low I/O locality, and large SSD capacity condition. When applying the SSD cache method with low locality and small SSD capacity, a number of data on the SSD cache are ejected. Since the data is written in the fast or slow tier, it is possible that the storage performance decreases in cases where write data is cached and the data is ejected from the SSD cache.

When applying the SSD cache method, the storage performance is better with F2-20% condition. From the result, it is considered that the SSD cache method is more effective with read-intensive, high I/O locality, and small SSD capacity. This is because the number of segments ejected from the SSD cache becomes smaller with high I/O locality workload, and the number of segments which need to be written in the fast or the slow tier also becomes smaller with read-intensive workload.

5.2 Result with SSD Tier Method and SSD Cache Method

Figure 6 shows the result with both methods that are used simultaneously and separately. The vertical axis shows the normalized storage I/O response time where the shorter storage I/O response time with the SSD tier method or the SSD cache method is 100%. This is the rate of the response time with the SSD tier method or the

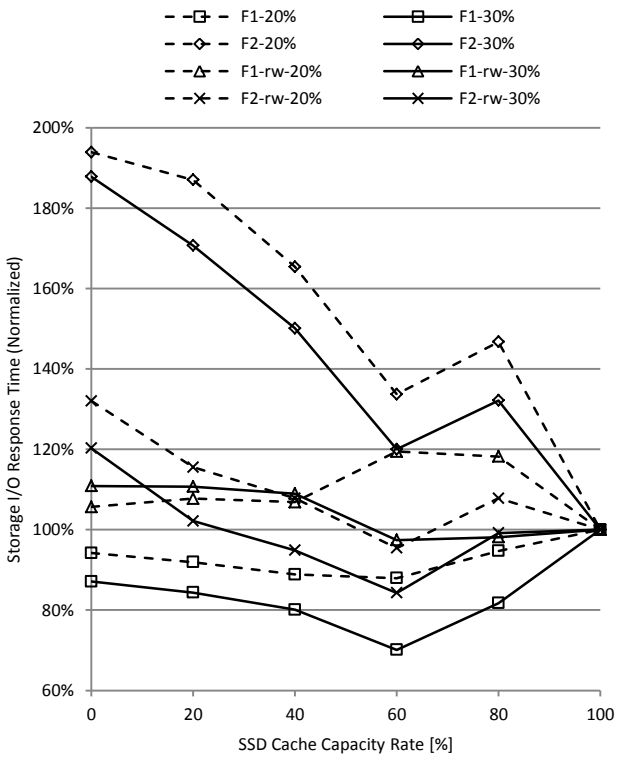


Figure 5. Storage I/O Response Time

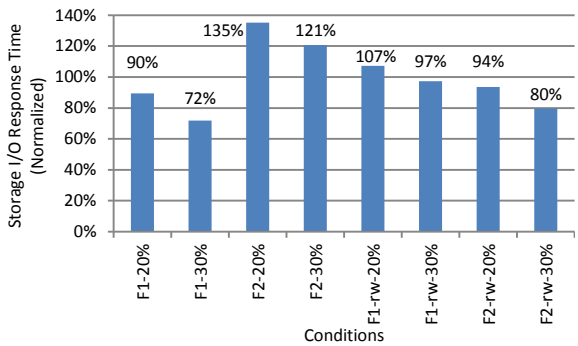


Figure 6. Storage I/O response time with both methods simultaneously

SSD cache method to the response time with both methods. The horizontal axis shows the experimental condition. When applying both methods simultaneously as well as the SSD tier method separately with large SSD capacity, the storage performance gets better.

With F1-rw-30% condition the difference in the storage performance between the conditions which both methods are used and the single method is used has a maximum gap and the storage I/O response time reduces by 28%. In F1 and F2-rw condition, the storage performance increases. From this result, it is considered that utilizing both methods simultaneously is effec-

tive with write-intensive workload. Compared to applying the SSD tier method, applying both methods is effective with not only write-intensive and low I/O locality workload but also write-intensive and high I/O locality workload.

6 Conclusion

We evaluated the storage I/O response time with the SSD tier method and the SSD cache method on the tiered storage which has SSD and HDD. The simulation result shows that the SSD tier method improves the storage I/O performance in write-intensive and low I/O locality workload with small SSD capacity. The SSD cache method is effective in read-intensive and high I/O locality workload with large SSD capacity. Utilizing both methods simultaneously with write-intensive workload is the most effective. The storage I/O response time is reduced by up to 28% utilizing both methods simultaneously compared to using the SSD tier method or the SSD cache method separately.

References

[1] J. D. Strunk: “Hybrid Aggregates: Combining SSDs and HDDs in a single storage pool”, ACM SIGOPS Operating Systems Review, Vol. 46, No. 3, pp. 50-56 (2012)

[2] F. Chen, D. A. Koufaty, and X. Zhang: “Hystor: Making the Best Use of Solid State Drives in High Performance Storage Systems”, Proceedings of the international conference on Supercomputing (ICS '11), pp. 22-32 (2011)

[3] S. Faibish, P. Bixby, J. Forecast, P. Armangau, and S. Pawar: “A new Approach to File System Cache Writeback of Application Data”, Proceedings of the 3rd Annual Haifa Experimental Systems Conference (SYSTOR '10), Vol. 3

[4] Hitachi Data Systems Corporation: “Dynamic Tiering”, (2013) http://www.hds.com/products/storage-software/hitachi-dynamic-tiering.html

[5] G. Schmidt, B. Dufrasne, J. Jamsek, P. Kimmel, H. Matsuno, F. Morais, L. Oxenham, A. Rainero, and D. Senin: “IBM System Storage DS8000 Easy Tier”, (2012) http://www.redbooks.ibm.com/redpapers/pdfs/redp4667.pdf

[6] EMC Corporation: “Fully Automated Storage Tiering for Virtual Pools (FAST VP)”, (2013) http://www.emc.com/storage/symmetrix-vmax/fast.htm

[7] Hewlett-Packard Development Company: “HP 3PAR Adaptive Optimization Software - Overview & Features”, (2013) http://h18006.www1.hp.com/storage/software/3par/aos/index.html

[8] S. Dawkins, K. Voruganti, and J. D. Strunk : “Systems Research and Innovation in Data ONTAP”, ACM SIGOPS Operating Systems Review, Vol. 46, No. 3, pp. 1-3 (2012)

[9] G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. Vogels: “Dynamo: Amazon's Highly Available Key-value Store” Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, Vol. 41, No. 6, pp. 205-220 (2007)

[10] B. F. Cooper, R. Ramakrishnan, U. Srivastava, A. Silberstein, P. Bohannon, H. Jacobsen, N. Puz, D. Weaver, and R. Yerneni: “PNUTS: Yahoo!'s hosted data serving platform” Proceedings of the VLDB Endowment, Vol. 1, No. 2, pp. 1277-1288 (2008)

[11] UMass Trace Repository: “OLTP Application I/O” (2007) http://traces.cs.umass.edu/index.php/Storage/Storage

Development and Experiment with Scheduling and Network Coding tools in Wireless Testbed

Seng Molika† **Kong Marry‡*** **Phal Des***

†Royal University of Phnom Penh
* Royal University of Phnom Penh
** Royal University of Phnom Penh

Abstract

In 802.11 wireless networks, when nodes are connected to wireless LAN, nodes are connected to an access point; therefore, nodes compete with each other for gaining access to the wireless medium. The competition causes error, contention problem, and thus there is problem with throughput that users could experience. In order to improve the performance throughput, logically, it might be good that the wireless medium is given priority to the long-delay-sensitive flow where sized-based scheduling, which it gives priority to the short flow, is adopted between users. Network Coding (NC) could also increase the bandwidth where the intermediate nodes are able to combine the packets coming from different flows, using Linear Network Coding, into one or several packets.

The goal of this paper is seeking for better performance throughput, where the experiment are implemented with both sized-based scheduling and NC techniques in a wireless testbed. The sized-based scheduling is implemented and compared to the standard scheduling applied on access point side. NC is implemented in both access point side and station side. Both techniques are implemented with Click Modular Router tool.

Keywords: Sized-based scheduling, Least Attained Service (LAS), Network Coding, Click

1 Introduction

In 802.11 wireless networks, when users connected to wireless LAN, users are connected to the same access point; therefore, users compete with each other to gain access to the wireless medium and losers might wait for a long period of time to gain access to the wireless medium; and so called contention problem. In wireless transmission, we also have problems of noise during the transmission.

To improve the performance and deal with the contention problem, one can use solution of scheduling. Scheduling is able to consider which users or which flow could access the medium. In addition to scheduling we also have Network Coding (NC).NC is in general used at intermediate nodes on the path between source and destination. Those intermediate nodes mix packets using linear network coding. Alternatively, we can use NC at a single location, the access point in our case, to combine several packets of several flows together.NC is also used at destination for decoding the combined packets in to original packets.

For better performance throughput, both techniques of Scheduling and Network Coding are studying in this paper. The objective of this paper is to implement both network coding and sized-based scheduling in a wireless testbed with Click Modular Router tool running on Linux fedora system.

2 Related Work

We begin with a summary of prior work on sized-based scheduling and network coding.

2.1 Sized-based Scheduling

In the network best effort model, all flows are transmitted without priority in term of flow size. On the contrary, network throughput as well as RTT could be improved by paying special attention to size of flows. Rai, Biersack, and Urvoy-Keller [1] and Biersack, Schroeder, and Urvoy-Keller [5] have studied sized-based scheduling policies which give priority to the short flows.

A typical example is Least Attained Service (LAS) where the next packet to be served is the

one that belong to the flow which received the least amount of service.LAS is also a buffer management mechanism. When the queue is full (unlike droptail) it inserts the appropriate packet in the queue and then drops the packet at the end of the queue. LAS intended to be deployed at the network edges where congestion can take place.

Implementation

$S_f = S_f + P$

Where f is *flow*, S_f is *accumulative value of service served for the flow*, and P is *packet size*. The lower the value of S_f, the higher the priority is.

The implementation with LAS scheduling gains advantages such as lower RTT during Slow Start phase and decrease of short flow losses. LAS also improves the mean response time since most TCP flows are short, while more than 50% of the bytes are carried by less than 5% of the largest flows for typical internet [1].

2.2 Network Coding

As mentioned above, in the 802.11 wireless network, when the AP transmits frames to one specific station in unicast mode, even though the other stations overhear the frames, they drop those frames since the MAC address did not belong to them. When NC is used, this behavior will be modified, so that stations store those frames for future use, even though the MAC address does not belong to the stations.

NC mechanism allows the intermediate nodes to combine several input packets into one or several output packets by using linear network coding. Some research papers such as XORs in the Air [6], and XOR Rescue [7], demonstrate the advantage of NC for gaining advantages in term of bandwidth and mean response times.

To improve the performance of network transmission, the COPE mechanism [6], which it uses network coding technique, can be used. The idea is described in figure 1 that after the router received frames from A and B, the router spends only one time slot for transmitting frames **a** and **b** to both stations. So, COPE spends only 3 times slots for the transmission.

With NC [8] that is used in COPE, after the

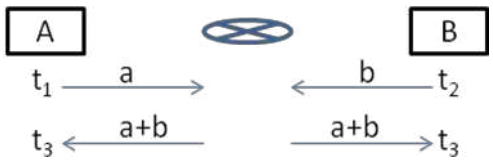


Figure 1. COPE

access point received all frames, **a** and **b**, with different destinations, it encodes the incoming frames into one encoded frame, a+b, then it sends out to the stations with the destination address of station A, for example. At the stations side, the promiscuous mode is set; thus, each station is able to receive the frames even though the destination address of the frames does not correspond to the stations. To get back the native frames from the encoded frames, the stations need to decode the encoded frame with the other native frames which are stored in their buffer pool and do not belong to the stations.

2.3 XOR Rescue

XOR Rescue (XORR) [7] is a NC-aided retransmission scheme which could improve the performance of the network transmission in term of bandwidth, expected goodput, and service time. XORR also keep overheard frames for future use as well.

Sender (AP) side

At sender side of XORR, AP has a queue for each station. For preventing queue reordering, only HOL from each queue could be the candidate for scheduling selection for doing the transmission. The queues are splitted into two groups, the Transmission queue (TxGroup) and Retransmission queue (RxQueue).

AP transmits frames to stations; and if the transmission to some stations fails, AP freezes the queues to those stations and places them to RxGroup. For the queues that have successfully received frames by the stations, AP still continues transmitting the frames until:

- The freeze time reaches the threshold **t**
- The Number of frozen queues reaches its threshold **thr**

Then, after **t** or **thr** is reached, the AP will do the retransmission from queues inRxGroup. If there are two or more frames to be retransmitted, the AP will do NC on some frames among the

frames to be retransmitted based on its coding metric. The XORed frames are called coded frames and the un-XORed frames are called native frames.

Receiver (station) side

The characteristic of XORR at receiver side is that when the receiver hears the native frames which belongs to it, the receiver transmits an acknowledgment (ACK) to the AP. If the receiver overhears the native frames that is not destined to it, those frames are stored in its native frame pool.

When the receiver hears or overhears the encoded frames, the receiver tries to decode the encoded frames immediately. The receiver can successfully decode the encoded frames only if it has other native frames in the native frame pool. After successfully decoding and if that native frames belong to the receiver itself, it will transmit the ACK to AP for removing the frames from queue at AP; however, if that native frames do not belong to the receiver, it will store the frames to its native frame pool. In contrast, if the receiver does not successfully decode the encoding frames, those frames are discarded.

2.4 Synergy between XOR Rescue and Sized-based Scheduling

The XOR Rescue mechanism, as described above, improves the network performance in the context of packet based opportunistic scheduling. Sized-based scheduling also improves the throughput of the transmission in the context of flow base scheduling. Thus, the goal of this project is to combine both techniques for better performance throughput. For the current time as a first step of this paper, the set up of a wireless testbed is made with such kind of transmission scheme.

The experiment is made on a wireless testbed as shown in figure 2 where a PC acts as the AP and the other two PCs are run as stations.

Assume that all devices, AP and stations, can hear each other well which mean that there is no hidden node problem; and the communication takes place only between AP and Stations.

With XORR mechanism and during the retransmission phase, the AP selects the frames to be encoded for maximizing the performance based on channel conditions. However, we

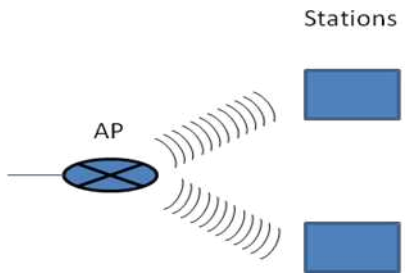


Figure 2. The design of first step in doing the experiment

decided not to implement this mechanism. Instead, the encoding operation is made with all retransmitted frames simultaneously when retransmission is triggered. The number of station is fixed. In our testbed, there are two stations.

3 Implementation

Sender Side

AP is installed with click where the flow chart of implementation is described in Figure 3. The configuration at AP is using XORR mechanism. In XORR mechanism, the click router receives frames from network interface cards eth0. Next, the AP implements encoding operation with LAS scheduling and then forwards frames to Wireless interface, wlan0.

Receiver Side

The station received encoded frames from AP, then the station do decoding operation to get back the native frames and after that it forwards the frames to upper layer and replies the ACK to AP. Figure 4is the receiver flow chart of XORR at the receiver (stations) side. The click router communicates with AP with wireless interface, wlan0.

4 Experimental Result

The performance of the network is studied and analyzed in this chapter since network performance is the very important point for users' requirement. Even though the theory shows that the new technologies are better than the old one, it is the main criteria of having the practical performance in order to support that new theory of the new technology.

Figure 5, 6, and 7 are performance comparison on TCP transaction between FIFO and LAS

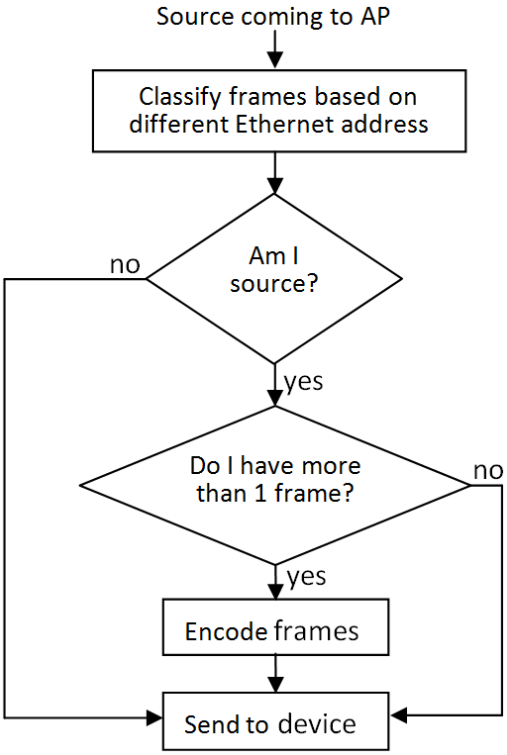


Figure 3. Flow chart of our implementation of sender (AP) side.

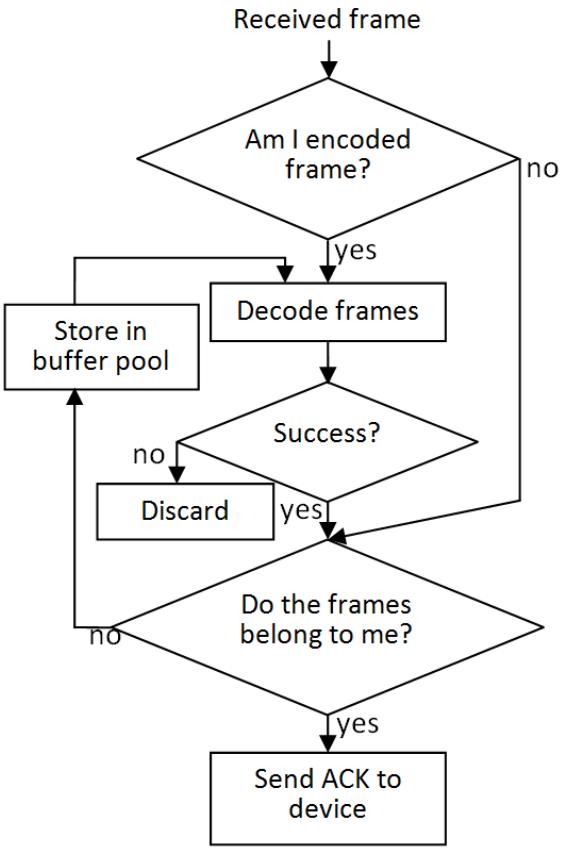


Figure 4. Flow chart of our implementation of receiver (AP) side

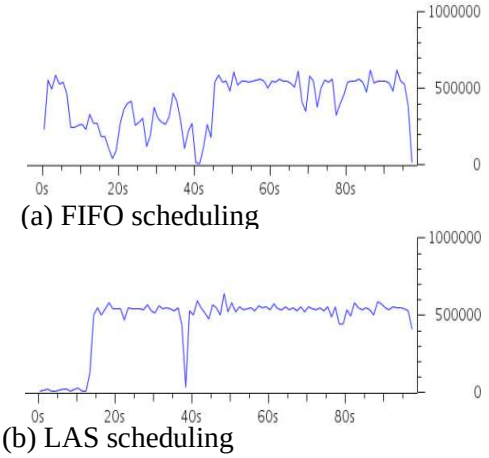


Figure5. FIFO and LAS scheduling performance throughput (x=time in second, y=bits)

scheduling which is analyzed with the bottleneck of 450 Kbit/s. The figures show the differences of network performance throughput between the access point implementing FIFO and LAS scheduling.

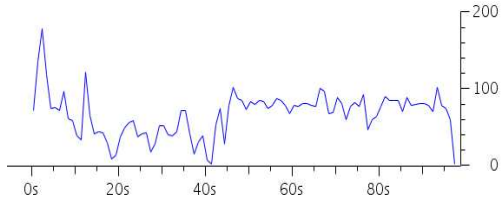
Figure 5 and Figure 6 describe real-time performance of Networking using FIFO and LAS in bits and in packets respectively. From the graphs, LAS scheduling is more stable than FIFO scheduling; as in Figure 5 (a) from 0s to 45s, the traffic of FIFO is vary from almost 0 bit to 550000 bits and the traffic of LAS in Figure 5 (b) is slow at the start, but after 15s the traffic is stable with rate of around 520000 bits. The graphs in Figure 6, describing the performance in packets, also have the same network performance as in bits in Figure 5.

Figure 7 shows the differences between re-transmission rates of FIFO scheduling and LAS scheduling in real-time performance. From the graph, the rate has range from 0 frames to 9 frames with retransmission rate of LAS a little higher than FIFO.

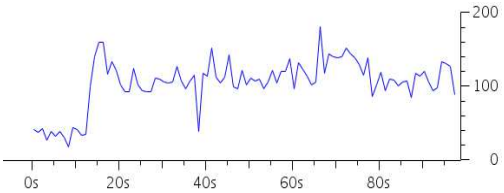
Even though the retransmission rate of LAS is higher than FIFO, but the difference is not much if compare to the rate of frames transmitted with real-time in figure 6. Therefore, in conclusion using LAS scheduling instead of FIFO scheduling implemented in an access point, users could have better network performance experience.

5 Conclusion

The two techniques of sized-based sche-duling and network coding run on click tool and

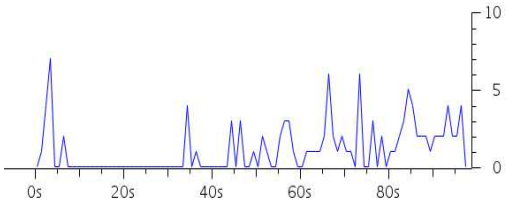


(a) FIFO scheduling

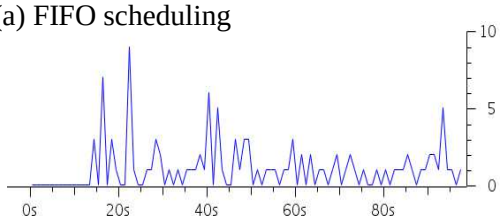


(b) LAS scheduling

Figure 6. FIFO and LAS scheduling performance throughput (x=time in second, y=packets)



(a) FIFO scheduling



(b) LAS scheduling

Figure 7. Frame Retransmission

using C++ language to build new elements could produce outcomes which the experiment could:

- Be implemented sized-based scheduling (LAS), FIFO scheduling and network coding (NC) on Click Modular Router tool.
- Obtain network analysis of sized-based scheduling (LAS) and FIFO scheduling.
- Obtain better performance throughput by implementing sized-based scheduling (LAS) compared to FIFO scheduling when they are implemented on an access point.
- Successfully transmitted and received by using Click Modular Router tool with NC implementation; and therefore, the better the transmission throughput after NC completion configuration.

For the full experiment of XORR for the project with performance analysis would be the future work.

Acknowledgment

I would like to express my appreciation and special thanks to my supervisors, Prof. Phal Des and Dr. Kong Marry. This is my great honor to be under their supervision.

I would like to thanks to lecturers at both universities, Royal University of Phnom Penh and University of Nice Sophia Antipolis, France where I am studying for my master degree. In addition, I thank professors and lecturers for providing me the very good knowledge. I am truly appreciated for their commitment, and responsibilities.

Last, I would like to thank my husband, Kep Samphy, and my family for they always encourages me to work harder for this paper.

References

[1] Sized-Based Scheduling to Improve the Performance of Short TCP Flows. I. A. Rai, E. W. Biersack, and G. Urvoy-Keller. 2005.

[2] Gast, Matthew. 802.11® Wireless Networks: The Definitive Guide. s.l. : O'Reilly, 2002. 0-596-00183-5.

[3] Convergence of propotional-fair sharing algorithms under general conditions. H. J. Kushner and P. A. Whiting. 2003, IEEE Trans. Wireless Communications.

[4] User-level performance of channel-aware scheduling algorithms in wireless data networks. S. C. Borst. 2005, IEEE Trans. Networking.

[5] Scheduling in Practice. E. W. Biersack, B. Schroeder, and G. Urvoy-Keller. 2007.

[6] XORs in The Air: Practical Wireless Network Coding. S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft. 2006.

[7] XOR Rescue: Exploiting Network Coding in Lossy Wireless Networks. F. Kuo, K. Tan, X. Y. Li, J. Zhang, and X. Fu. 2009, in the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON).

[8] Network Coding: An Instant Primer. C. Fragouli, J. Y. Le Boudec, and J. Widmer. 2006.

[9] The Click modular router. R. Morris, E. Kohler, J. Jannotti, and M. F. Kaashoek. 1999.

[10] http://www.read.cs.ucla.edu/click/

[11] Driver- Linux Wireless. [En ligne] http://wireless.kernel.org/en/users/Drivers/
[12] http://linuxwireless.org/en/users/Documentation/hostapd
[13] http://docs.fedoraproject.org/en-US/Fedora/15/html/Deployment_Guide/s1-dhcp-configuring-server.html

Noise Suppressor using Zero Phase Signal and Accumulated Spectrum Technique

Weerawut Thanhikam†
†Faculty of Engineering and Technology
Panyapiwat Institute of Management, Thailand
weerawuttha@pim.ac.th

Abstract

This paper proposes a wide-band noise reduction method using temporal accumulated spectrum and zero phase (ZP) signal. In the previous study, we replace the ZP signal around the origin with the ZP signal in the second or latter period to get an estimated speech ZP signal. For very low SNR environment, reliable period estimation is difficult. This paper presents a study of period estimation in noisy speech based on temporal-spectral representation. Since harmonic structure of speech changes much more slowly than noise spectrum, spectral peaks related to period harmonics would stand out over the noise through the accumulation. In this paper, we combine robust period estimation technique and ZP noise suppressor. The major advantages of this method are that it can reduce not only stationary wide-band noises but also non-stationary wide-band noises and does not require a prior estimation of the noise spectral amplitude. Simulation results show that the proposed noise reduction method further improves the SNR especially in a low SNR environment.

Keywords: zero phase signal, temporal accumulated, noise reduction

1 Introduction

In recent years, noise reduction technique is required in a wide area of applications including mobile communication and speech recognition systems. The major examples are cell-phone and car navigation systems, since they are widely used in various noisy environments. Although many noise reduction methods have been proposed in decades [1]–[3], noise reduction still remains as an annoying problem. Especially, wide-band noise reduction is one of the most important issues. In a practical environment, two kinds of wide-band noise exist. The first one is a stationary wide-band noise which is generated from air-conditioner, PC fan, moving car in a tunnel, driers and so on. The other is a non-stationary wide-band noise which is generated from thunder, clap, and other bangs. Examples of wide-band noise are depicted in Fig. 1, where (a) shows a female speech signal corrupted with stationary white noise, and (b) shows one corrupted with non-stationary impulse noise. They were sampled at 8kHz. It is well known that both of a white noise and an impulse noise have a flat spectral amplitude (i.e., all spectral amplitudes have the same value). From this fact, we see that they belong to wide-band noise, while their time domain properties are obviously different. Because of such different properties, stationary and non-stationary noise reduction methods have been individually studied.

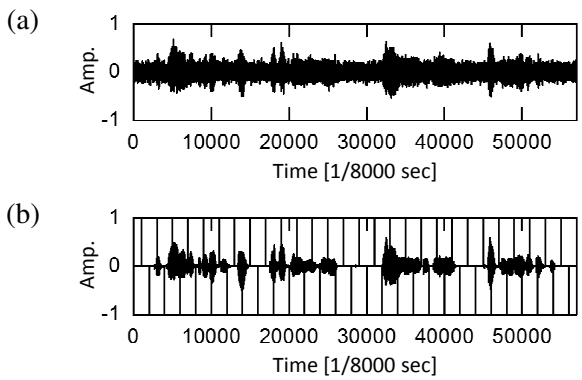


Figure 1. Examples of speech signal mixed with wide-band noise: (a) speech corrupted with white noise, (b) speech corrupted with impulse noise.

The spectral subtraction [1] is the most popular among stationary noise reduction techniques. it has been extensively researched. The spectral subtraction has been evolved to more efficient methods so-called spectral gain approaches [4]–[7]. However, those methods also require a priori estimated noise spectral amplitude. Thus, they

cannot reduce non-stationary noise such as an impulse noise. On the other hand, the median filter [2] is often used to remove an impulse noise. As a result, it achieves to remove impulse noises and keep edge components of the input signal, simultaneously. Unfortunately, the median filter is ineffective to reduce a stationary noise.

The goal of this study is to improve a wide-band noise reduction system which utilizes a zero phase (ZP) signal which is defined as the IDFT of a spectral amplitude [8]. As shown in [8], a white noise and an impulse noise can be reduced by processing the ZP signal only at the origin. However, this method is not applicable for other noises. To improve effectiveness of noise reduction, we extend the concept of [8] and pitch period estimation. In this paper, we combine period estimation proposed in [11] with the ZP noise suppressor proposed in [12] for suppressing noise which includes both of stationary and impulsive noise in very low SNR evironment. Simulation results show that the proposed method is effective to reduce such wide-band noises.

This paper is organized as follows. In Section 2, we briefly describe spectral gain method for stationary noise reduction. In Section 3, we introduce the ZP signal and explain its properties. In Section 4, we derive a noise reduction method for reducing wide-band noises. Section 5 gives experimental results to confirm the effectiveness of the proposed method.

2 Spectral Gain Method for Stationary noise reduction

In this section, we present an overview of the spectral gain method for stationary noise reduction, As we will see in Section 4, this system is a foundation of the proposed wide-band noise reduction method.

A general spectral gain method is shown in Fig. 2. Here, $x(n)$ is an observed noisy signal at time n, and it composed of a clean speech $s(n)$ and an additional noise $d(n)$ given as

$$x(n) = s(n) + d(n). \tag{1}$$

The noisy signal $x(n)$ is transformed into frequency domain by segmentation and windowing with a window function $h(n)$, e.g., Hanning window. The DFT coefficient of the noisy signal at

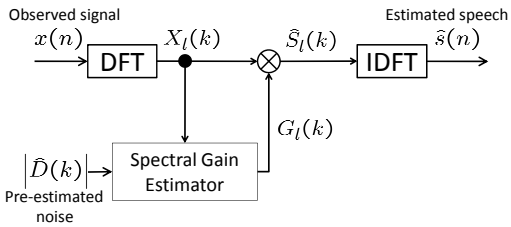


Figure 2. Overview of general spectral gain method.

frame l and frequency bin k is calculated with

$$X_l(k) = \sum_{n=0}^{N-1} x(lQ+n)h(n)e^{-j2\pi nk/N}, \tag{2}$$

where N denotes the DFT frame size. The window is shifted by Q samples for the computation of the next DFT. The DFT coefficient $X_l(k)$ also consists of speech and noise parts, as given by

$$X_l(k) = S_l(k) + D_l(k), \tag{3}$$

where $S_l(k)$ and $D_l(k)$ represent the DFT coefficients obtained from $s(n)$ and $d(n)$, respectively. As shown in Fig. 2, the speech enhancer calculates a speech spectral gain $G_l(k)$ by using a priori estimated noise spectral amplitude $|\hat{D}(k)|$, where $|\cdot|$ denotes the spectral amplitude. Here, various definitions of $G_l(k)$ have been proposed [4]–[7], e.g., the spectral subtraction's spectral gain is $G_l(k) = 1 - |\hat{D}(k)|/|X_l(k)|$. The enhanced speech spectrum $\hat{S}_l(k)$ is given by

$$\hat{S}_l(k) = G_l(k)X_l(k). \tag{4}$$

Finally, we obtain the enhanced speech $\hat{s}(n)$ in time domain after the IDFT of $\hat{S}_l(k)$ and overlap-add.

To achieve non-stationary noise reduction, we have to estimate the speech spectral gain $G_l(k)$ without a priori estimation of noise spectral amplitudes. One of the solutions is obtained by mapping the observed signal into the ZP domain.

3 Zero Phase Signal

In this section, we explain about the ZP signal [8]. By using the ZP signal, we will reveal a wide-band noise reduction method in Section 4.

For simplicity, we omit the frame index l. The DFT coefficient of $x(n)$ can be expressed as

$$X(k) = |X(k)|e^{j\angle X(k)}, \tag{5}$$

where $\angle\{\cdot\}$ denotes the spectral phase. We define the ZP signal of $x(n)$, $x_0(n)$, as

$$x_0(n) = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^{\beta} e^{j\frac{2\pi n}{N}k}, \qquad (6)$$

where β is a certain constant. Obviously, $|X(k)|^{\beta}$ can be reproduced from the DFT of the ZP signal $x_0(n)$ as

$$|X(k)|^{\beta} = \sum_{n=0}^{N-1} x_0(n) e^{-j\frac{2\pi k}{N}n}. \qquad (7)$$

Through this paper, we put $\beta = 1$. In addition, we assume that $x(n)$ is a real valued signal. In this case, the ZP signal $x_0(n)$ come to real even signals [8].

Here, we show a few examples of the ZP signal. Let the spectral amplitude $|X(k)|$ be a constant α_0 (≥ 0). Substituting $|X(k)| = \alpha_0$ into (6) with $\beta = 1$, we have

$$x_0(n) = \alpha_0 \delta(n), \qquad (8)$$

where $\delta(n)$ denotes the Kronecker's delta function. Eq.(8) shows that the ZP signal of any flat spectral amplitude is expressed as the delta function. From derivation in [8], we have

$$x_0(n) = \sum_{m=1}^{\lfloor \frac{N}{2k_c} \rfloor} \frac{\alpha_m}{N} \cos\frac{2\pi m k_c}{N} n, \qquad (9)$$

where $\lfloor\cdot\rfloor$ denotes a floor function, and α_m is an amplitude of the mth frequency. Hence, the ZP signal of a periodic signal becomes also a periodic signal whose period is N/k_c.

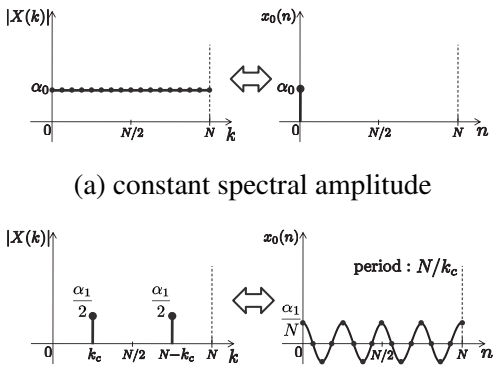
4 Wide-Band Noise Reduction Using Zero Phase Signal

4.1 Derivation of Noise Reduction System

As the same manner of conventional noise reduction methods [4]–[7], we also assume that the spectral phase of the estimated speech signal is equal to that of the observed signal, i.e., $\angle\hat{S}(k) = \angle X(k)$. It means that

$$x_0(n) = s_0(n) + d_0(n), \qquad (10)$$

where $s_0(n)$ and $d_0(n)$ are the ZP signals of $s(n)$ and $d(n)$, respectively. Under the assumption, we derive a wide-band noise reduction system



(a) constant spectral amplitude

(b) equally spaced line spectra

Figure 3. Examples of zero phase signal: (a) constant, (b) equally spaced line spectra.

which can reduce both of stationary and non-stationary wide-band noises, without a priori estimation of noise spectral amplitudes.

Firstly, we models a speech signal $s(n)$ in a short observation as a HNM (Harmonic plus Noise Model) [9], [10] given as

$$s(n) = \sum_{m=1}^{\lfloor \frac{N}{2k_c} \rfloor} \alpha_m \cos(2\pi \frac{k_c}{N} mn + \theta_m) + \varepsilon(n), \qquad (11)$$

where k_c/N is the normalized fundamental frequency, and α_m and θ_m are the amplitude and the phase of the mth harmonic frequency, respectively. The signal $\varepsilon(n)$ is a noise signal generated by passing a white noise through an all-pole filter [10]. Here, we assume that the energy of $\varepsilon(n)$ in an observation frame is sufficiently small in comparison to one of the harmonic part. This assumption is appropriate for a voiced speech, but it is not appropriate for an unvoiced speech. Although this assumption may give a degradation to an enhanced speech, the degradation is not fatal. Because, voiced speech energy is usually much greater than unvoiced one.

Next, we show the properties of practical noise and speech signals in the ZP domain. We plotted the ZP signals of some practical wide-band noises and a female speech signal in Fig. 4, where (a) shows a tunnel noise, (b) shows a motor noise, (c) shows a babble noise, (d) shows a clap noise, (e) and (f) show voiced and unvoiced speech signals, respectively. Here, all signals were sampled at 8kHz and $N = 256$. We see from Figs. 4(a)–(d) that the energy of all wide-band noises is concentrated around time 0 in the

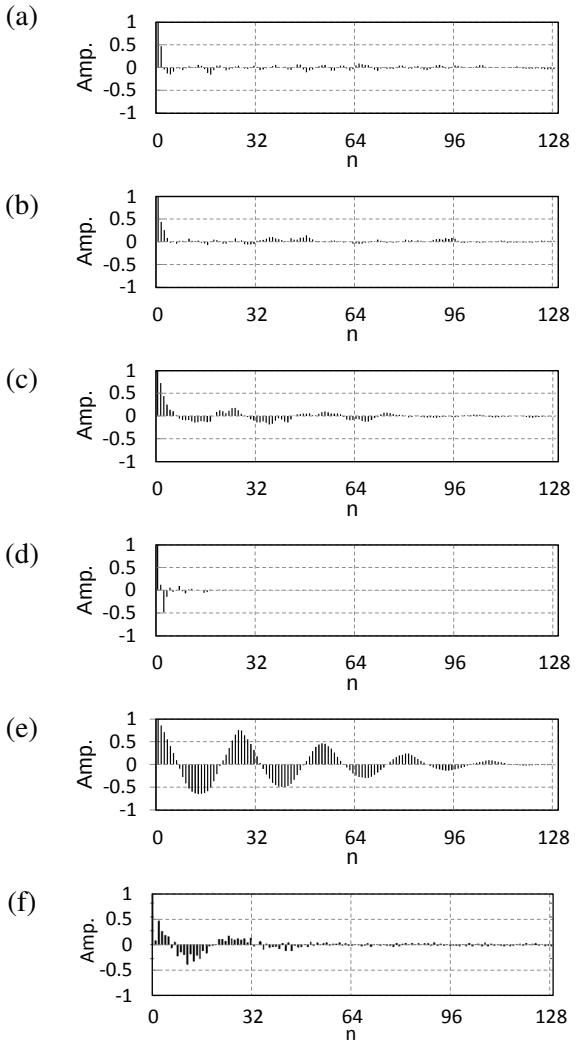


Figure 4. Zero phase signals. (a) tunnel noise, (b) motor noise, (c) babble noise, (d) clap noise, (e) voiced speech signal, (f) unvoiced speech signal.

ZP domain. Hence, if we remove the ZP signal around the origin, then the noise is greatly reduced. On the other hand, from Fig. 4(e), we see that the voiced speech becomes a periodic signal with amplitude attenuation in the ZP domain. This attenuation arises due to the window function. Since the window function is known, we can compensate the attenuation. We also see from Fig. 4(e) that the effect of $\varepsilon(n)$ is extremely low for the voiced speech. On the other hand, the ZP signal of the unvoiced speech shown in Fig. 4(f) is similar to that of the noises. As shown in Fig. 4(e) and (f), the energy of the unvoiced speech is less than the voiced one. In this paper, we concentrate on extracting the voiced speech

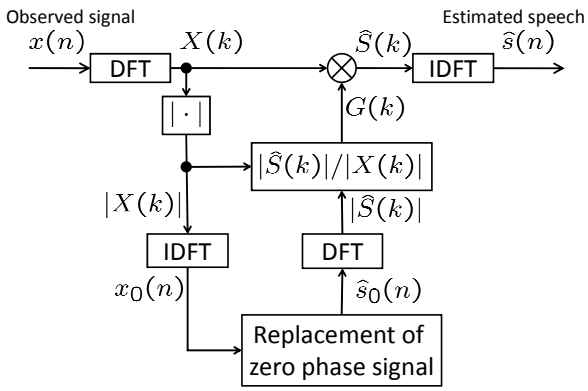


Figure 5. Proposed wide-band noise reduction system using zero phase signal.

rather than the unvoiced one.

The noise ZP signal has nonzero values mainly around origin. Hence, we assume that the noise ZP signal $d_0(n)$ at $(n > L)$ is sufficiently small for $x_0(n)$. Then we have

$$\begin{aligned} x_0(n) &\approx \begin{cases} s_0(n) + d_0(n), & 0 \le n \le L \\ s_0(n), & L < n \le \frac{N}{2}, \end{cases} \\ x_0(n) &= x_0(N-n), \quad \frac{N}{2} < n < N. \end{aligned} \tag{12}$$

When the pitch period of the speech ZP signal, $T = N/k_c$, is greater than L, we can estimate T as the time index of the second peak of $x_0(n)$. Since the observed ZP signal $x_0(n)$ in $T \le n < N + L$ does not include the noise components, we obtain the estimated speech ZP signal $\hat{s}_0(n)$ by the following replacement.

$$\hat{s}_0(n) = \begin{cases} \mathrm{sc}(n) \cdot x_0(T+n), & 0 \le n \le L \\ x_0(n), & L < n \le \frac{N}{2} \end{cases}, \tag{13}$$

where $\mathrm{sc}(n)$ is a scaling function to compensate the envelope attenuation of the speech ZP signal. It is obtained as the reciprocal function of the window for signal segmentation. When we use the hanning window, the scaling function $\mathrm{sc}(n)$ is given as [12]

$$\mathrm{sc}(n) = \frac{1 + \cos\frac{2\pi}{N}n}{1 + \cos\frac{2\pi}{N}(n+T)}. \tag{14}$$

After the replacement (13), the DFT of $\hat{s}_0(n)$ gives the estimated speech spectral amplitude $|\hat{S}(k)|$. Finally, taking the IDFT of $|\hat{S}(k)|e^{j\angle X(k)}$, we have the estimated speech signal $\hat{s}(n)$ in time domain.

Fig. 5 shows the block diagram of the proposed wide-band noise reduction system, where the spectral gain is given as $G(k) = |\hat{S}(k)|/|X(k)|$. Here, this system requires the additional DFT and IDFT to achieve stationary and non-stationary wide-band noise reduction without a priori estimation of noise spectral amplitudes. One of most important parameters in the proposed method is the replacement size L which is presented in [12]. Another most important one is the pitch period T. We proposed to accumulate zero phase spectral over consecutive frames. In the next section, we investigate about an estimation method of them with proposed technique.

4.2 Temporal Accumulated Zero Phase Spectrum

Let $\mathbf{s}_0^{(l)} = [s_0^{(l)}(0)s_0^{(l)}(1) \cdots s_0^{(l)}(\frac{N}{2})]$ denote the zero phase spectrum of the *l*th frame of speech signal. In voiced speech, the zero phase spectrum retains the harmonic peaks that can be used for period estimation. However, there exist many irrelevant spectral peaks, especially when the signal is corrupted by noise. To make the harmonic peaks more prominent, we propose the temporally accumulated zero phase spectrum (TAZPS), which is defined as

$$\mathbf{p}_0^{(l)} = \mathbf{s}_0^{(l-\lfloor\frac{K}{2}\rfloor)} + \cdots + \mathbf{s}_0^{(l)} + \cdots + \mathbf{s}_0^{(l-\lfloor\frac{K}{2}\rfloor+K-1)}, \tag{15}$$

where K is the number of accumulated frames. Note that the more use of K causes delay in the estimation process. Since the speech has a quasi-stationary characteristic, its harmonic structures of ZP speech signal in neighboring frames are expected to change slowly. Pitch period of ZP signal is very likely to occur at the similar time index in the previous and next frames. Therefore, the temporal accumulation as shown in (15) would reveal peak of ZP signal at the pitch period. On the other hand, peak of ZP signal caused by noise likely to be more randomly distributed along time axis, such that their locations are unpredictable in the accumulated spectrum.

Figure 6 shows an example of computing the temporally accumulated ZP spectrum. Zero phase spectra of three consecutive frames in a clean speech signal are shown in Fig. 6(a), (b) and (c). The TAZPS in Fig. 6(d) is obained by accumulating the three ZP spectra, i.e., $\mathbf{p}_0^{(l)} = \mathbf{s}_0^{(l-1)} + \mathbf{s}_0^{(l)} + \mathbf{s}_0^{(l+1)}$, It can be seen that the harmonic structrues are well retained in the ZP spectra as well as in the TAZPS. When the same speech segment is corrupted by additive noise, the harmonic structures in the single become less predictable as shown in Fig. 7(a), (b) and (c). In the TAZPS as shown in in Fig. 7(d), prominent peaks are concentrate around the pitch period. Pitch period estimation can be done by searching the second peaks of ZP signal. In the next section, we investigate about an estimation of pitch period from TAZPS.

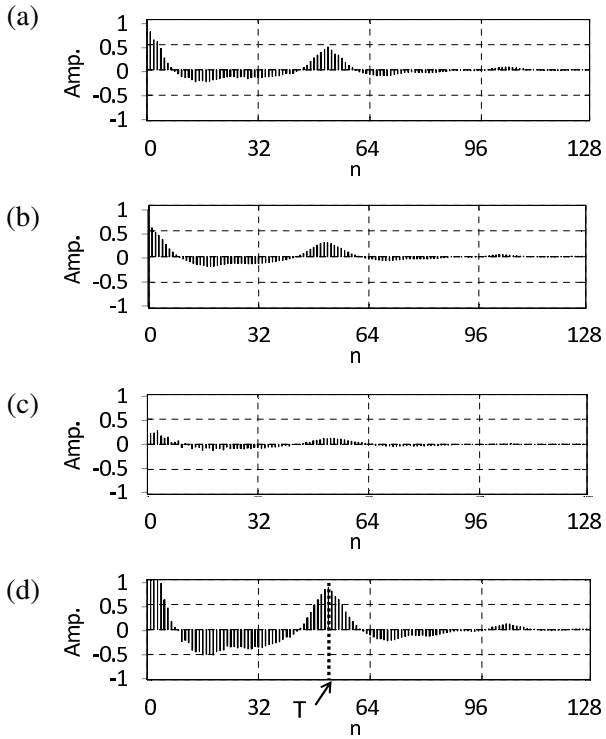


Figure 6. Zero phase signals. (a) $\mathbf{s}_0^{(l-1)}$, (b) $\mathbf{s}_0^{(l)}$, (c) $\mathbf{s}_0^{(l+1)}$, and (d) temporally accumulated zero phase spectrum (TAZPS).

4.3 Parameter Estimation

We first describe about an estimation method of the pitch period T, and then discuss an appropriate replacement size L in an empirical manner.

From the definition (6), we see that any ZP signal takes the maximum value at the origin. On the other hand, as shown in Fig. 6, a voiced speech provides a periodic ZP signal with amplitude attenuation. However, as we stated in the previous section, TAZPS improved robustness against noisy ZP signal, the TAZPS will be

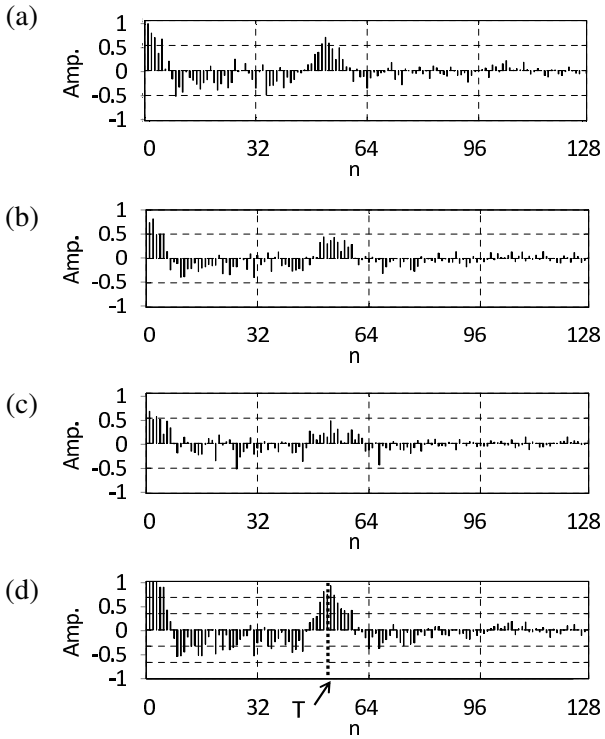


Figure 7. Zero phase signals of speech added with white noise at -10 dB SNR. (a) $\mathbf{x}_0^{(l-1)}$, (b) $\mathbf{x}_0^{(l)}$, (c) $\mathbf{x}_0^{(l+1)}$, and (d) temporally accumulated zero phase spectrum (TAZPS).

used to find T instead of using single-frame ZP signal. As reported in [14], an averaged pitch period of male speakers is about 8ms, and that of female speakers is about 4ms. Hence, an computationally efficient peak search method can be established by restricting the search range. The pitch period T is given as

$$T = \arg\max_{t_L \leq n \leq t_H} \{\mathbf{p}_0(n)\}, \tag{16}$$

where, t_L is the lowest index number of the search range, and t_H is the highest one.

Next, we choose the replacement size L in an empirical manner as proposed in [12]. The results for the four practical noises with Input SNR of 0dB are shown in [12]. We employ $L = 10$ as an appropriate value for wide-band noise reduction, especially reducing the non-stationary clap noise. In the next section, we perform noise reduction simulations to confirm the effectiveness of the TAZPS method with $L = 10$.

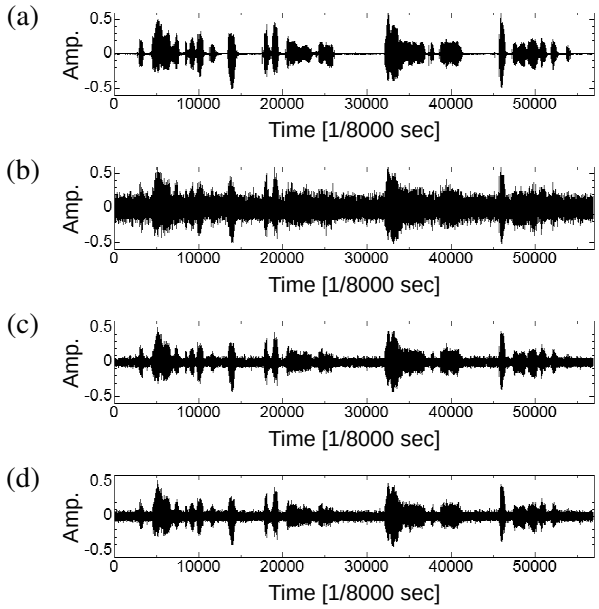


Figure 8. Results of white noise reduction: (a) clean speech signal, (b) speech signal corrupted by white noise (SNR=0.0dB), (c) the estimated speech by the spectral subtraction method (SNR=7.0dB), (d) the estimated speech by the proposed method (SNR=6.8dB).

5 Simulation

We evaluate the capability of the proposed method in further detail. The speech signals used in the simulations were taken from ATR-promotion database [13]. All signals used in simulations were sampled at 8kHz. We put $N = 256$ and $L = 10$, and used the Hanning window for signal segmentation. We put $t_L = 16$ and $t_H = 64$ that implies the pitch search range from 2ms to 8ms. The proposed method was compared with some conventional methods.

Firstly, we performed the noise reduction simulations for input SNR of 0dB, where we used artificially generated white and impulse noises. In the simulations, we compared the proposed method with the most traditional spectral subtraction method [1]. The waveforms of the white noise reduction results are shown in Figs. 8(a)–(d), where (a) shows the clean speech signal, (b) shows the speech signal corrupted by the white noise with the input SNR of 0.0dB, (c) shows the estimated speech signal obtained by the spectral subtraction method, where the output SNR was 7.0dB, and (d) shows the estimated speech signal obtained by the proposed method, where the out-

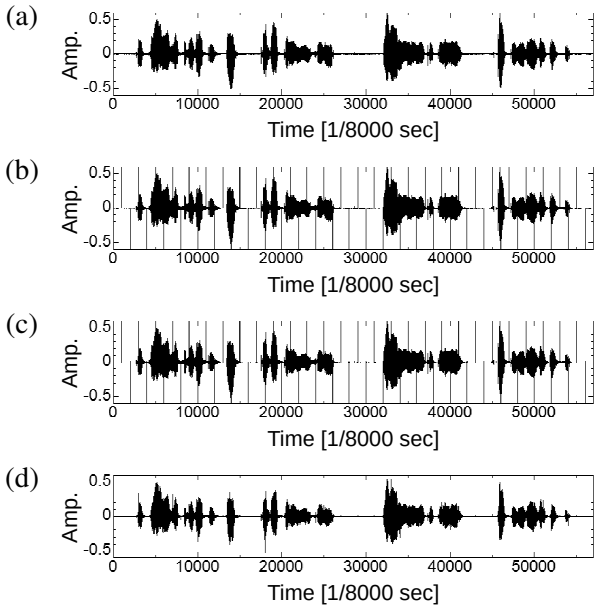


Figure 9. Results of impulse noise reduction: (a) clean speech signal, (b) speech signal corrupted by impulse noise (SNR=0.0dB), (c) the estimated speech by the spectral subtraction method (SNR=-0.1dB), (d) the estimated speech by the proposed method (SNR=13.3dB).

put SNR was 6.8dB. From these results, we see that the proposed method can reduce the stationary wide-band noise, without a prior estimation of noise spectral amplitudes. For the impulse noise reduction simulation, we imposed the condition that amount of impulse per block-segment is one or zero on this simulation. The results are summarized in Figs. 9(a)–(d), where (a) shows the clean speech signal, (b) shows the speech signal corrupted by the impulse noise with the input SNR of 0.0dB, (c) shows the estimated speech signal by the spectral subtraction method, where the output SNR was -0.1dB, and (d) shows the estimated speech signal by the proposed method, where the output SNR was 13.3dB. From these results, we see that the proposed method can suppress the non-stationary impulse noise with the same procedure of stationary wide-band noise reduction. The spectral subtraction method cannot reduce such non-stationary wide-band noise.

Table 1 show the output SNR of the wide-band noise reduction results. We see from the results for the stationary wide-band noise shown that the noise reduction capability of the proposed method is almost the same or slightly low in comparison to ones of SS and VMAP methods which require the prior estimation of the noise spectral amplitude. On the other hand, CZPS and the proposed method do not require any prior estimation of the noise spectral amplitude. In the stationary practical noise cases the proposed method is superior to the CZPS. This improvement is caused by removing the noise ZP signal $d_0(n)(1 \leq n \leq L)$ which cannot be reduced by the CZPS method. Moreover, TAZPS also give more accurate pitch period estimation in proposed method. Note that the capability of the proposed method exactly reaches to ones of SS or VMAP when we utilize the prior estimation of the noise spectral amplitude. On the other hand, for non-stationary wide-band noises i.e., clap noise, the noise reduction capability of the proposed method and CZPS are superior to SS and VMAP. When the input SNR was 0dB in clap noise situation, the proposed method improved the SNR to 11.6dB which is almost 6dB higher than the result of the CZPS.

6 Conclusion

In this paper, we have proposed a noise suppressor using zero phase signal and accumulated peak spectrum technique. The noise reduction is achieved by replacing the observed ZP signal around the origin with the ZP signal in the estimated period using TAZPS. The proposed method does not require a prior estimation of noise spectral amplitudes, and can reduce not only stationary wide-band noises but also non-stationary wide-band noises. Simulation results have shown the effectiveness of the proposed noise reduction method. The stationary wide-band noise reduction capability of the proposed method is almost the same or slightly low in comparison to the spectral subtraction method and the variable MAP method which require a priori estimation of noise spectral amplitudes. The most advantage of the proposed method is that it can provide a high noise reduction performance for non-stationary wide-band noises. In the clap noise reduction, the proposed method attained the output SNR of 11.6dB when the input SNR was 0dB. The effectiveness of the proposed method for the other wide-band noises was also confirmed. Future works include to derive an extraction method of unvoiced speech in the ZP domain.

Table 1. Output SNR of wide band noise reduction results [dB]

System	-10.0 [dB]			-5.0 [dB]			0.0 [dB]		
	W	C	WI	W	C	WI	W	C	WI
SS [1]	−2.0	−9.9	−5.1	2.4	−4.9	−0.3	6.9	0.1	4.5
VMAP [5]	0.4	−9.9	−5.3	−4.8	−4.9	1.2	9.4	0.1	6.3
CZPS [8]	−2.0	−3.6	−0.1	2.5	1.4	4.3	7.2	6.1	8.6
Proposed method	−1.3	9.9	0.2	2.8	11.0	4.16	6.6	11.6	7.6

W : white noise C : clap noise WI : white and impulse noise

References

[1] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-27, no.2, pp.113–120, April 1979.

[2] M. Muneyasu and A. Taguchi, Nonlinear digital signal processing, Asakura Publishing Company, Tokyo, 1999.

[3] A. Kawamura, Y. Iiguni and Y. Itoh, "A noise reduction method based on linear prediction with variable step-size," IEICE Trans. Fundamentals, vol.E88-A, no.4, pp.855–861, April 2005.

[4] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-32, no.6, pp.1109–1121, Dec. 1984.

[5] Y. Tsukamoto, A. Kawamura and Y. Iiguni, "Speech enhancement based on MAP estimation using a variable speech distribution," IEICE Trans. Fundamentals, vol.E90-A, no.8, pp.1587–1593, Aug. 2007.

[6] A. Kawamura, W. Thanhikam, and Y. Iiguni, "A speech spectral estimator using adaptive speech probability density function," Proc. of EUSIPCO 2010, pp.1549–1552, Aug. 2010.

[7] W. Thanhikam, A. Kawamura and Y. Iiguni, "Speech enhancement using speech model parameters refined by two-step technique" Proc. of the Second APSIPA Annual Summit and Conference, p.11, Dec. 2010.

[8] Y. Kamamori, A. Kawamura, and Y. Iiguni, "Zero phase signal analysis and its application to noise reduction," IEICE Trans. Fundamentals, vol.J93-A, no.10, pp.658–666, Oct. 2010.

[9] D.W. Griffin and J.S. "Lim. Multiband-excitation vocoder," IEEE trans. Acoust., Speech, Signal Processing, ASSP-36, pp. 1223–1235, Aug. 1988.

[10] J. Laroche, Y. Stylianou, and E. Moulines,"HNS: Speech modification based on a harmonic + noise model" in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing '93, Minneapolis, MN, pp. 550–553, Apr. 1993.

[11] F. Huang and T. Lee, "Pitch estimation in noisy speech based on temporal accumulation of spectrum peaks," in *Proc. Eur. Signal Process. Conf.*, Aug. 2011, pp. 451-455.

[12] W. Thanhikam, A. Kawamura, Y. Iiguni, "Stationary and non-stationary wide-band noise reduction using zero phase signal" IEICE Trans. Fundamentals, Vol.E95-A, No.5, pp.843-852, May. 2012.

[13] "SLDB, Natural Speech and Language Database," ATR-Promotions, http://www.atr-p.com

[14] S. Furui, Digital speech processing, Tokai University Press, Tokyo, 1985.

The assignment of chunk size according to the target data characteristics in deduplication backup system

Mikito Ogata[†] **Norihisa Komoda**[‡]
[†]Hitachi Information and Telecommunication Engineering, Ltd.
781 Sakai, Nakai-machi, Ashigarakami-gun, Kanagawa, 259-0180
{mikito.ogata.mz@hitachi.com}
[‡]Osaka University
2-1, Yamadaoka, Suita, Osaka, 565-0871
{komoda@ist.osaka-u.ac.jp}

Abstract

This paper focuses on the trade-off between the deduplication rate and the processing penalty in backup system which uses a conventional variable chunking method. The trade-off is a non-linear negative correlation if the chunk size is fixed. In order to analyze quantitatively the trade-off all over the factors, a simulation approach is taken and clarifies the several correlations among chunk sizes, densities and average lengths of the different parts. Then it clarifies to assign an appropriate chunk size based on the data characteristics dynamically is effective to weaken the trade-off and provide higher efficiency than a conventional way.

Keywords: Deduplication, Backup, Archive, Capacity Optimization, Enterprise Storage

1 Introduction

Due to the explosive increase of the data in IT system, backup operation are becoming a burden to the system because of the resource usage, the processing time and the managing cost, while it is an indispensable operation to recover the data in case of unpredictable disaster. Major requirements to backup operations are to shorten the processing time and to reduce the resource usage, especially storage capacity to keep the data for a long term. Recently, the technology called deduplication has become popular to reduce the burden. This is a technology to eliminate the duplication of the backup target data and store only unique data in the storage. The reduction of stored data reduces not only the backup storage cost but also other resources running workload.

Various techniques have been proposed so far to provide more deduction with less processing time from the point view of more cost-effective backup operation[1][2]. However, because the deduplication rate and the processing time have a non-linear negative correlation, it is difficult to improve both rate and time simultaneously if only one chunking size is activated. This paper shows how to assign the appropriate chunking size according to the data characteristics. The assignment weakens the trade-off in single chunking method and therefore provides more efficient deduplication than any conventional approach.

2 Deduplication backup system

2.1 Variable chunking algorithm

In usual customers environment, backup target data include various types of files and how much the duplication patterns remains or how they locates depends on the customers environments or applications, for example some are scarcely duplicated because it was much edited, changed or updated, some are densely duplicated because it was rarely changed, just replicated or copied. In this paper, all changed or updated parts of data are called ”differed area” and unchanged parts are called ”identical area”.

A deduplication backup system has a processing module which eliminates the duplication within the target data. A module reads the target data, divides into many small segments, distinguishes the duplicated area from unique or newly updated area, then transfers and stores only the unique data in the storage. Figure 1 shows the typical operational flow of the deduplication process from reading the target data and finalizing to store the unique data. The size of the box is not to scale.

A module divides the target data in small segments called ”chunk” which are the unit of reduction or storing (Chunking). Two dividing al-

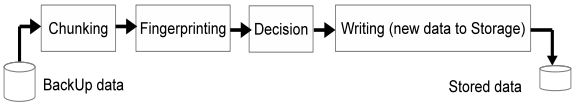


Figure 1. Deduplication process.

gorithms are widely implemented, fixed length chunking or variable length chunking. Fixed length chunking divides the data into chunks with the same length that is defined in advance. Variable length chunking divides into chunks with variable length that is calculated by the data patterns. A typical algorithm of variable length chunking scans the data from the beginning of the file one byte at a time to the end by shifting a fixed length window[3][4]. The algorithm generates a special value, called a signature, based on the window using a Rabin's algorithm[5]. Then when the signature is recognized to match a predefined value, called "anchor", the end byte of the windows is set at the end of the chunk. Otherwise the window is shifted by one byte and generate a signature again. The average chunk size is defined from how many bits are taken as a signature. For example, the length of the signature is 12 bits, under the assumption that the calculation of signature generates completely random values, 2^{12} patterns are possible, this results in 4KB chunk size in average.

Next, the module calculates the unique code from the chunk data to analyze the similarity of chunks using. The hashing algorithm is commonly used to calculate the code for example by SHA-1, SHA-256. This code is called by 'fingerprint' (Fingerprinting).

Next the module decides the uniqueness of each chunk using the fingerprint (Decision), whether the chunk is "duplicated" that the same chunk has been stored or "not duplicated" that all previous chunks are different. Finally, the module write out only "not duplicated" chunks into the storage (Writing).

2.2 Issues

Many approaches have been proposed and implemented in order to increase the efficiency of deduplication. For the customers, two criteria are important to discuss the efficiency, one is the deduplication rate and the other is the processing time. The deduplication rate is desirable to be higher and the processing time is desirable to be shorter. How to choose the adequate chunk size of the algorithm is one key sensitive factor, because that affects heavily to the efficiency. In many cases, the chunk size is determined based on the consideration to keep the balance of the performance and the deduplication capability for an assuming hypothetical environment[6][7].

In order to gain more reduction, to assign smaller chunk size is generally effective. A smaller chunk size can pick up the differed area in more precise than a bigger chunk size. However, the effectiveness of the chunk size, that is how much sharply the chunk size cut off the duplication, very much depends on the distribution of differed area among the data, such as a length of each differed area and a distance between differed areas. When a small chunk size is assigned to longer differed data, many continuous chunks may be wasted to remain not duplicated area. When differed area are located with similar distance to the chunk size, many chunks may include both of differed area and identical area, this results in poor deduplication rate. When differed area are distributed with much smaller or much bigger distance comparing to the chunk size, the effectiveness is improved.

Processing time, that is how much time or resource utilization is necessary to reduce the duplication, is also depends on the chunk size and others. A smaller chunk size requires more CPU and IO intensive processing time, a bigger one does less time. The less duplicated, less storing operation to the storage. Further, the correlation of the process time and the chunk sizes are non-linear, for example, for a smaller chunk size, the processing time increases more steeply as the size decreases. This non-linearity makes the compensation of trade-off hard to overcome.

In any conventional way, only one chunk size is assigned for all data or all environments, it is not effective. If the chunk size is assigned to gain better deduplication rate with smaller processing time corresponding to the characteristics of data or environmental, it is beneficial for the system.

3 Simulation

3.1 Implementation

The followings are the operation and assumptions of our simulator. As a convenience, we call the simulator "SIM".

(1) Backup data
Target backup data is implemented as a list with 1024*1024 length, one entity represents 64 bytes length of real data. Therefore, it represents in total 64MB real data length. In the stream, there are many area that represents "identical", identical to one of previously stored area and many area which represents "differed", not identical to anyone of previously stored area. SIM reduces all of identical area as much as it can and remains definitely all of differed area by the process.

The ratio of the amount of all identical area to the total amount of data is defined r as a "duplication rate". The domain of r is $0 \leq r \leq 1$. In other words, the ratio of all differed area comes to $1 - r$. These differed areas are spread with equal density over the stream.

(2) Chunking
SIM assumes an anchor algorithm to separate the data into chunks. The signature length to control the average chunk size is set to 11 to 15 bits. These come to 2 to 32 KB in average chunk size respectively. In this case, the sizes of chunks fits to geometrical distribution[8]. The distribution parameter p is equal to $1/2^n$, here n is the window size in bits. We call the case that the average chunk size is m as a notation of (m) for convenience. From industrial practical reason, SIM set two boundaries of the minimum and the maximum size. The minimum size is set to the half of chunk size, and the maximum size is set to one and half of the chunk size. For example, when an average chunk size is 8KB, the minimum size is 4KB and the maximum size is 12KB. Table. 1 is a list of parameters of the simulation.

Table 1. Parameters of simulation

Parameter	Values
Duplication rate: r	0.2, 0.3, 0.4, ... 0.9
Length of differed area	2, 4, 8, ... 64 [KB]
Average chunk size: m	2, 3, 4, ... 64 [KB]

(3) Decision
SIM checks whether the chunk includes differed area. When the chunk includes no differed area at all, SIM decides the chunk should be removed. When the chunk includes fully or partially differed area, SIM decided the chunk should be remained. The "deduplication rate" is calculated as a ratio of total amount of chunks which are decided to be removed to total amount of data. In addition, the fulfillment rate is calculated as a ratio of the deduplication rate to the duplication rate. The "number of reduced chunks" is counted as a number of identical therefore removed chunks by SIM. Since the processing time of deduplication is very varied on the resource configuration, usage in the system, SIM evaluates the total number of removed chunks as a uniform substitute.

(4) Environment
SIM is coded by the language "R 2.15.1 GUI 1.52 64-bit", operated on the environment of the hardware consists of Intel Core i7 2.6GB CPU, 8GB (1600MHz, DDR3) memory.

3.2 Evaluation of the validity of SIM comparing the measured data

At first, the validity of SIM is evaluated by comparing the equivalent data that are gathered from the measurement using real hardware. Our experiment hardware is based on a linux based NAS (Network Attached Storage) product connecting to HDD storage, NFS (Network File System) interface and a variable chunking mechanism is coded. Chunk size is controlled as an parameter of average chunk size for NAS and controlled as $\frac{1}{p}$ for SIM. These parameters are called here "operational chunk sizes". Total amount of data in measurement is 1GB, the length of differed area is fixed at 32KB. The differed areas are randomly distributed over the data with control parameters of duplication rate that are assumed to reflect practical data characteristics. NAS reports the deduplication rate, the average chunk size, the number of reduced chunks by the experiment.

Table 2 shows the comparison of two chunk sizes gathered from each experiment. One is chunk size that NAS divides in the measurement, the other is one that SIM divides in the simulation. Simulated chunk sizes match well to measured chunk sizes in the area of smaller operational chunk size. Simulated chunk size become gradually shorter than measured as operational chunk size increases. This comes from the difference that NAS uses multiple chunking algorithm installed, on the other hand, SIM uses only one variable chunking algorithm.

Table 3 and figure 2 show the comparison of deduplication rates. The rates matches in smaller chunk size, 2% difference in case of 8KB. The

Table 3. Comparison of simulated and measured deduplication rate

Operational chunk size	measured				simulated				difference [%]			
	0.50	0.66	0.80	0.95	0.50	0.66	0.80	0.95	0.50	0.66	0.80	0.95
2KB	0.470	0.634	0.782	0.945	0.469	0.638	0.792	0.948	-0.3	0.7	1.2	0.3
4KB	0.449	0.614	0.768	0.940	0.441	0.615	0.777	0.943	-1.9	0.2	1.2	0.3
8KB	0.403	0.573	0.738	0.931	0.393	0.573	0.754	0.930	-2.6	0.0	2.1	-0.1
16KB	0.319	0.494	0.680	0.913	0.305	0.522	0.708	0.921	-4.4	5.7	4.0	0.9
32KB	0.200	0.373	0.584	0.880	0.190	0.391	0.631	0.896	-4.9	4.6	8.2	1.8
64KB	0.083	0.231	0.452	0.829	0.080	0.248	0.500	0.849	-4.0	7.6	10.8	2.4

Table 2. Comparison of simulated and measured chunk size

Operational chunk size	measured [KB]	simulated [KB]	difference [%]
2KB	1.81	1.81	0.0
4KB	3.58	3.69	-3.1
8KB	7.06	7.31	-3.4
16KB	13.9	14.69	-5.4
32KB	26.9	29.38	-8.4
64KB	52.4	58.75	-10.7

difference become larger as the chunk size increases. The rate become worse for NAS as the duplication rate increases. This comes from the difference that in NAS some additional information is embedded in to the stream to improve the data integrity, SIM does not. The embedded information spoil the deduplication efficiency and lessen the rate as the chunk size increases. The rate is better for NAS in case that the duplication rate equals to 0.5. This comes from the difference that NAS uses multiple chunking algorithms, results in higher deduplication rate, at the same time NAS has the penalty to lessen the rate by embedded information which becomes bigger as duplication increases.

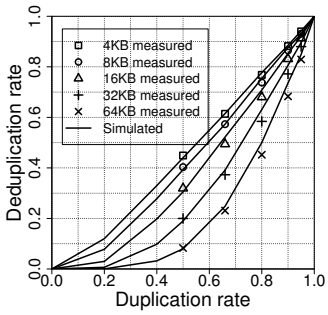


Figure 2. Deduplication rate by Single layer deduplication system

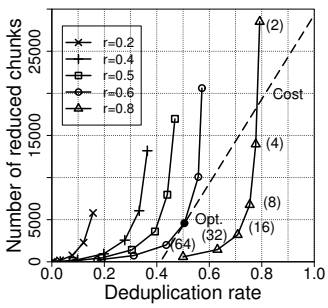


Figure 3. Reduced number of chunks by duplication rate (length of differed area = 32KB)

4 Evaluation

4.1 Dependency on duplication rate

Figure 3 shows the correlation of deduplication rate and the number of reduced chunks by changing the duplication rate of 0.2, 0.4, 0.5, 0.6, 0.8 under the same length of differed size of 32KB. The number of reduced chunks that is removed by SIM increases as the duplication rate increases or the chunk size decreases. A criteria is useful to analyze the effects of changing chunk sizes. As shown in the figure, as the chunk size decrease starting from 32KB, the number of reduced chunks becomes 3.2 times by (16) from (32), the deduplication rate improves 60% from 0.190 to 0.305. Then, the number of reduced chunks becomes 2.5 times by (8) from (16), the deduplication rate improves 30%, successively 2.0 times and only 3% by (2) from (4).

The improvement ratio decrease as chunk size decreases. This means to decrement chunk size in bigger chunking situation gains more improvement with less penalty. At the same time, the improvement is more clear as the duplication rate decrease.

The deduplication rate and the fulfillment rate with each chunk size increase as the duplication rate increase. This comes because the deduplication works more efficiently as the potential of mixture of identical area and differed area in the same chunk becomes lower along duplication rate increase.

4.2 Dependency on length of differed area

Figure 4 shows the correlation between the deduplication rate and the number of reduced chunks by changing the length of differed area of 2, 4, 8, 16, 32KB under the same duplication rate of 0.6. As the length of differed area increase, the corresponding number of reduced chunks increase and deduplication rate also increase under the same chunk size. This comes from the decrement of deduplication inefficiency because the potential of mixture of identical area and differed area in a chunk become lower as length of differed area increase.

The improvement of the deduplication rate along decreasing of chunk size becomes smaller as the length of differed area increases. This comes because the preciseness of reduction by using small chunks degrades.

The number of reduced chunks increase and therefore the deduplication rate increase as the length of differed area increases for all chunk sizes. The improvement becomes smaller for small chunk sizes. For example, when the length of differed area increase from 2KB to 32KB, the improvement of deduplication rate is 1.8 times in case of (2), 3.3 times for (4), 9.5 times for (8), and so on. This indicates to assign bigger chunk size is more effective than the smaller ones in case that the length of differed area widely varies in the data.

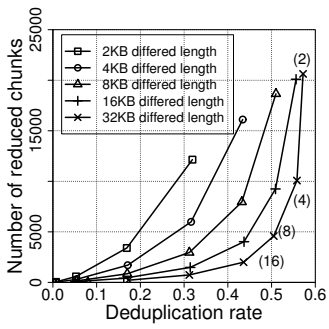


Figure 4. Reduced number of chunks by length of dirty (duplication rate = 0.6)

4.3 Assignment of optimal chunk size

As described, there is a trade-off between the number of reduced chunk size and the deduplication rate. Therefore, when the problem to assign the optimal chunk size is defined as a linear programming with two objectives, to maximize the deduplication rate and to minimize the number of reduced chunks, it generates a Pareto-optimal solutions, because there exists no feasible solution for which an improvement in one objective does not lead to a simultaneous degradation in the other objectives. In figure 3 or figure 4, each solid line represents one corresponding Pareto-optimal solutions. Generally, to assign the best chunk size as an optimal one depend on the customers choice, considering of various factors from their proprietary priorities[9]. Without neglecting practical priorities, this paper simplify the customers factors and priorities in two costs. One factor is the operational cost, which includes the cost of resources like server, LAN(Local Area Network)、SAN(Storage Area network), utilities and the cost of management. The other factor is the capital cost like Storage and the cost of maintenance. The priorities is to minimize the total cost. Here, the cost is defined in the following function, the optimal chunk size is assigned to minimize the cost.

$C = A * ReducedChunks - B * DedupedRate$

Here,

$ReducedChunks$: Number of reduced chunks

$DedupedRate$: Deduplication rate

Figure 3 shows an example of cost function.

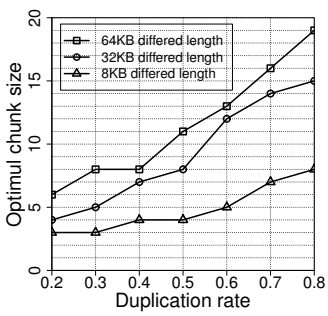


Figure 5. Optimal chunk size to minimize the cost (length of dirty =8、32、64KB)

The dotted line indicates the cost for corresponding deduplication rate and number of reduced chunks, A is set to 1 and B to 50000. In case of duplication rate of 0.6, the chunk size to pro-

vide the minimum cost is 8KB.

Figure 5 shows the chunk size to minimize the cost. The optimal chunk size increase as the duplication rate increases. This comes from the characteristic of bigger chunks that reduce coarsely with short process time. Further, the optimal chunk size increase as the length of differed area increases. This comes from the same reason. In addition, the difference between the optimal chunk size increases as the duplication rate increases. This comes because the effectiveness to use bigger chunks becomes weak as the duplication rate increases in case of short length of differed area.

In commercially available products, 4KB chunk size is a popular size. Table 4 shows the improvement for the cases of duplication rate of 0.2, 0.4, 0.6, 0.8 and the length of differed area of 8, 32, 64KB. The improvement of optimal chunks increases as the duplication rate and the length of differed area increase, becomes 33% in the case of 0.8 and 64KB.

Table 4. Improvement of cost value by the optimal chunk size to conventional 4KB.

Length of differed area		Duplication rate			
		0.2	0.4	0.6	0.8
8KB	Optimal	(3)	(4)	(5)	(8)
	Improve	27.4	0.0	5.0	13.8
32KB	Optimal	(4)	(7)	(12)	(15)
	Improve	0.0	12.6	18.1	30.4
64KB	Optimal	(6)	(8)	(13)	(19)
	Improve	4.4	16.5	29.7	33.6

5 Conclusion

The paper simulates variable chunking algorithm in backup deduplication system. First, the simulation clarify changing chunk size is more beneficial for bigger chunks than smaller chunks, gains more improvement with less penalty. In addition, the improvement is increase as the duplication rate decrease. Next, to assign bigger chunk size is more effective than the smaller ones in case that the length of differed area widely varies within the data. Finally, under the reasonable assumption to minimize the cost, the optimal chunk size is 3KB in case of 0.2 duplication rate and 8KB length of differed area with 27% improvement than fixed 4K chunk size, similarly the optimal chunk size is 19KB in case of 0.8, 64KB with the improvement of 33%.

References

[1] Y. Tan et al. Dam: A data ownership-aware multi-layered de-duplication scheme. In *2010 Fifth IEEE International Conference on Networking, Architecture and Storage*. IDC-Japan, 2010.

[2] Y. Won, J. Ban, J. Min, L. Hur, S. Oh, and J. Lee. Efficient index lookup for de-duplication backup system. In *Modeling, Analysis and Simulation of Computers and Telecommunication Systems, 2008. MASCOTS 2008. IEEE International Symposium on (Poster Presentation)*, Sept. 2008.

[3] U. Manber. Finding similar files in a large file system. In *Proceedings of the USENIX Winter 1994 Technical Conference*, 1994.

[4] A. Muthitacharoen, B. Chen, and D. Mazières. A low-bandwidth network file system. In *Proceeding of SIGOPS. 18th Symposium on Operating Systems Principles.*, Banff, Canada, 2001.

[5] M. O. Rabin. Fingerprinting by random polynomials. Technical report, Department of Computer Science, Harvard University, 1981.

[6] D. Meister and A. Brinkmann. Multi-level comparison of data deduplication in a backup scenario. In *Proceedings of SYSTOR 2009, The 2nd Annual International Systems and Storage Conference*. ACM, May 2009.

[7] G. Wallace, F. Douglis, H. Qian, P. Shilane, S. Smaldone, M. Chamness, and W. Hsu. Characteristics of backup workloads in production systems. In *Proceedings of the 10th USENIX Conference on File and Storage Technologies*, 2012.

[8] J. Min, D. Yoon, and Y. Won. Efficient deduplication techniques for modern backup operation. In *IEEE Transaction on computer, Vol. 60, No. 6*, 2011.

[9] M. Ogata and N. Komoda. Optimized assignment of deduplication backup methods using integer programming. In *Proceedings of JCIS2011: The 4th Japan-China Joint Symposium on Information Systems*, Apr. 2011.

Optimizing Acceleration-based Activity Models with User's Physical Characteristics Information

Yousuke Yasuda† **Takuya Maekawa**†
†Graduate School of Information Science and Technology, Osaka University
{yasuda.yosuke,maekawa}@ist.osaka-u.ac.jp

Abstract

In the activity recognition task using body-worn accelerometers, we can construct an activity recognition model with placing small burdens on an end user by employing the crowd-sourcing approach, i.e., obtaining training data from numerous numbers of other users. However, the training data are noisy and decrease recognition accuracy for the end user's training data. In this study, we attempt to find training data from other users that decrease the recognition accuracy with user's physical characteristics information to train an activity model adapted to the end user.

Keywords: Acceleration Sensor, Daily Activity Recognition

1 Introduction

Recently small and cheap sensors have been proliferated due to the advance in sensing technologies, and thus various sensors are embedded in our daily life environment. In addition, there are many studies on activity recognition technology to recognize human daily activity such as "walk," "run," and "sleep" by using body-worn accelerometers. Activity recognition technology has various kinds of real-world applications such as care of the elderly, fitness support, and lifelogging.

Activity recognition technology is a technology that estimates what a user is performing by using sensors attached to the user's body or installed in the environment. This paper focuses on the body-worn acceleration sensor based approach. Because supervised machine learning technique is used to achieve this goal, an end user should prepare labeled training data by himself. However, it is not practical from the viewpoint of cost and burden. Therefore, a lot of activity recognition studies try to reduce the cost and burden. For example, some recent studies have attempted to collect labeled training data from the crowds, i.e., employing participatory sensing approach. That is, these studies assume that the unspecified number of users collect training data.

Generally, learning a high quality activity model requires a large quantity of training data because the training data can cover various activity patterns. However, training an activity model using training data collected by the crowd reduces activity recognition accuracy for an end user's sensor data because sensor data from other users who have different physical characteristics from the end user are also different from those obtained from the end user.

Assume that we train an activity model of a right-handed user. It is clear that sensor data related to "brushing teeth" activity of a right-handed user are greatly different from those of a left-handed user. In such case, we consider that, when we learn a right-handed user's model, we should discard "brushing teeth" training data obtained from left-handed users. This is because the training data from left-handed users can degrade the learned model if the training data are similar to those related to other activities of the right-handed user. That is, because training data from the crowd include data from users whose physical characteristics are very different from those of the end user of interest, we should discard such confusing and unnecessary training data. In this study, we find and discard such confusing training data automatically by using user's physical characteristics information.

2 Related Work

We introduce activity recognition studies that attempt to reduce the labeling effort required of an end user. Ohmura et al.[2] recognize activities based on small amounts of end user sensor data by using adaptation techniques such as maximum-likelihood linear regression (MLLR) and maximum a posteriori (MAP) adaptation techniques, which are usually used in speech

recognition. Forster et al.[3] implement an adaptive gesture recognition system that employs brain decoded signals to detect recognition errors automatically, and re-train the recognition model according to the detected error. Stikic et al.[4] reduce labeling costs and achieves highly accurate recognition by using an activity learning technique. That is, the recognition system locates a sensor data segment that it finds hard to recognize, and then asks the end user to input the correct answer for the segment. Huynh et al.[5] also reduce labeling effort by combining generative models and discriminative classifiers. More specifically, they first obtain clusters of activities by employing a generative approach without any labels and and then boost the recognition results of the generative models by employing a discriminative approach with small numbers of labels. Krassnig et al.[6] model the activities of an end user by using labeled sensor data from other users of same gender as the end user. In this study, we construct activity models adapted the end user and improve activity recognition accuracy by filtering the training data thought to reduce activity recognition accuracy using the physical characteristics information of the end user.

3 Proposed method

This study tries to improve activity recognition accuracy by removing training data that are confusing and may not be observed from an end user by using physical characteristics information.

3.1 Overview of proposed method

Our proposed method consists of the four following procedures.

1. Feature extraction
2. Identifying mixed area
3. Finding unnecessary training data
4. Learning activity models

The overview of the proposed method is shown in Fig. 1.

Firstly, the proposed method extracts features from sensor data. In this study, we assume time-series data obtained from body-worn accelerometers. Therefore we extract feature for each sliding time window and construct a feature vector

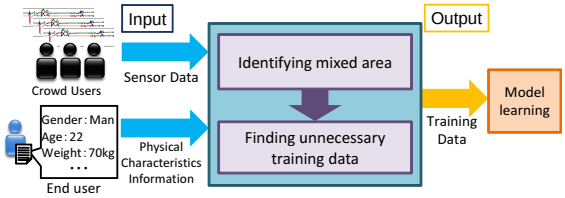


Figure 1. Overview of the proposed method

for each window. By using feature vectors from the crowd, we find areas in the feature space where feature vectors belonging to different activity classes are mixed.

To find a mixed area exists between two activity classes, we first model distribution for each class with Gaussian Mixture Model (GMM) and then find an area where these distributions are overlapped with each other. The mixed area decreases the activity recognition accuracy because we can not obtain a hyperplane that separates these classes well. To cope with the problem, we remove data points in the area belonging to one class of the two classes.

3.2 Feature extraction

We extract features from the crowd user's sensor data. We extract features based on existing activity recognition studies. Because we assume time-series acceleration data, we compute a feature vector for each sliding time window. We extract features based on the Fast Fourier Transform (FFT) components of 64 sample time windows. As features, we use the mean, energy, and dominant frequency, as described below. The mean is the DC component of the FFT coefficients, and can characterize the posture of parts of the body. For example, a mean corresponding to hand posture during tooth brushing may have particular characteristics. The energy feature is calculated by summing the magnitudes of the squared discrete FFT component. Note that the DC component of the FFT coefficients is excluded from this summation. The energy can be used to distinguish low intensity activities such as standing from high intensity activities such as walking. The dominant frequency is the frequency that has the largest FFT component, and it allows us to distinguish between repetitive motions with similar energy values. We construct a feature vector concatenating the above features extracted from all the body-worn accelerometers.

3.3 Identifying mixed area

3.3.1 Coping with imbalanced data

Before modeling the training data, we should cope with the imbalanced data problem. In many cases, one activity class is represented by a large number of data samples while the other is represented by only few. Such data is called imbalanced data [8] and it has negative effect on the activity recognition task. In this study, classes represented by large data sets are down-sized so as to match the size of the fewest one.

3.3.2 Learning activity model using GMM

We model sensor data distribution for each class by using Gaussian Mixture Model (GMM). A GMM consists of several gaussian distributions and is mathematically represented as formula (1). Note that x is a random variable vector, μ is a mean vector of multidimensional gaussian distribution, Σ is a variance-covariance matrix, and λ_m is a mixture ratio of gaussian distribution. Also, formula (2) shows the n dimensional gaussian distribution.

$$p(x) = \sum_{m=1}^{M} \lambda_m N(x; \mu_m, \Sigma_m) \quad (1)$$

$$N(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} exp - \frac{(x-\mu)^T}{\Sigma(x-\mu)} \quad (2)$$

$$(\sum_{m=1}^{M} \lambda_m = 1) \quad (3)$$

When we learn parameters of a GMM, we use the Expectation-Maximization (EM) algorithm . The EM algorithm is an algorithm that maximizes the expectation of the likelihood function by repeating Expectation Step (E step) and Maximization Step (M step).

3.3.3 Finding mixed area

By using GMMs created from sensor data from the crowd, we find an area where two GMMs (two activity classes) are overlapped with each other. Such area is referred to be a mixed area. We explain a technique to identify training data points included in the mixture area. We can compute a likelihood value of a training feature vector for each activity model by using formula (1). When a vector has the highest likelihood value for an activity class that is different from that the vector actually belongs to, we assume that the vector is in the mixed area.

3.4 Finding unnecessary training data

3.4.1 Obtaining criteria to decide unnecessary data

From the data in the mixture area, we want to obtain a criteria, “a set of data points belonging to activity A is unnecessary for end users with a certain physical characteristics, and another set of data points belonging to activity B are unnecessary for other users”. A labeled training data point has physical characteristics information of a user in addition to an activity information. Because we know a data point is obtained from which user, we also know physical characteristics information related to the data point. Therefore, for each mixture area, we construct a the decision tree whose input is physical characteristics information and output is an activity class, and it becomes a criteria that determines which data point is necessary for which user.

Fig. 2 shows decision trees where data points are divided according to gender or dominant hand information. Data points in a root node show those in a mixed area, and data points are classified into one of two leaf nodes according to its physical characteristics information. As shown in the right tree, the data points are separated with higher purity. From the right example, we can say that activity A usually happens for a right-handed user, and action B usually happens for a left-handed user in the mixed area. In other words, sensor data of activity B are unnecessary for a right-handed user, and sensor data of activity A are unnecessary for a left-handed user. We use C4.5 algorithm for to make decision tree. The example of a decision tree made from the mixed area between “Vacuuming” and “Drawing Whiteboard” is shown in Fig. 3 .

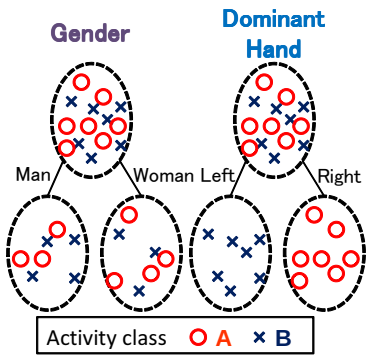


Figure 2. Example decision trees for mixed area

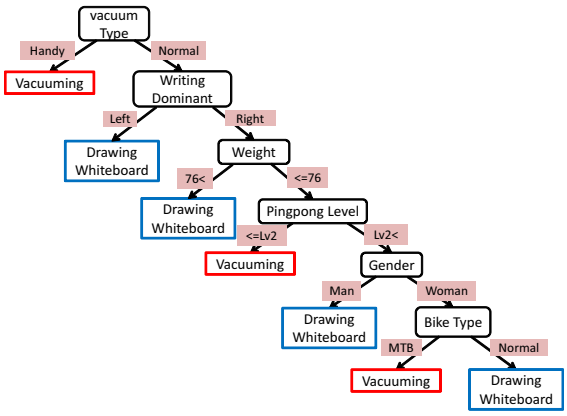


Figure 3. Decision tree of mixed area between "Drawing Whiteboard" and "Vacuuming"

3.4.2 Identifying unnecessary training data

A mixed area is an area where the training data of two kinds of activity classes are mixed. In other words, the decision tree made from the training data in the mixed area is a binary classifier as shown in Fig. 3 (classifies data points into "Drawing Whiteboard" or "Vacuuming" class). By inputing the physical characteristics information of an end user, we can know which class of activity is more likely to happen for the end user in the mixed area. That is, we can determine which data point is used to train the end user's activity recognition model.

3.5 Training activity model

After filtering training data, we learn an activity model of an end user. We use C4.5 decision tree whose input is feature vector and output is an activity class. We recognize unlabeled test data obtained from an end user by using the decision tree.

4 Evaluation experiment and discussion

4.1 Dataset

We collected sensor data with our developed wireless sensor nodes equipped with three-axis acceleration sensors and sampling rates of $30Hz$. Each participant wore the sensor nodes on the wrists of both hands, waist, and right thigh. Here, the most natural data would be acquired from the normal daily lives of the participants. However, obtaining sufficient samples of such data from large numbers participants is very costly. Therefore, we collected sensor data by using a semi-naturalistic collection protocol that permits greater variability in participant behavior than laboratory data. In the protocol, the participants perform a random sequence of activities following instructions on a worksheet. The participants are relatively free as regards how they perform each activity because the instructions on the worksheet are not very strict, e.g., "brush your teeth at the sink" and "vacuum the room with a hand-held vacuum cleaner." During the experimental period, the participants completed data collection session that included the random sequence of activities listed in Table 1. These activity were basically selected from those reported in acceleration-based activity recognition studies. Each participant completed ten sessions in total in our experimental environment. The sensor data from the four sensor nodes attached to the participant were also sent to the mobile computer. We collected labeled activities from the 60 participants.

Each participant also filled out a questionnaire that asked for the PC information listed in Table 2. We selected various kinds of PC information about the participants ranging from basic physical characteristics information such as weight and gender to information related to the activities listed in Table 1. The activity-related information includes the dominant hand used in several activities, and the frequencies of several activities in their daily lives. The PC information also includes the types of objects used in the experiment, i.e., type of bicycle and type of vacuum cleaner.

Table 1. List of daily activities

Standing	Brushing Teeth
Walking	Washing Dishes
Running	Using PC
Sitting	Drawing Whiteboard
Ascending Stairs	Writing Notebook
Descending Stairs	Play Pingpong
Bicycling	Vacuuming

Table 2. Physical Characteristics Information

Age	Chair Type
Height	Writing Experience
Weight	Pingpong Experience
Gender	Capability Blindtouch
Pingpong Dominant	Dishwash Frequency
Writing Dominant	Bike Frequency
Brushing Dominant	Vacuum Type
Vacuuming Dominant	Pingpong Level
Bike Type	

4.2 Evaluation methodology

We evaluated our recognition method with "leave-one-participant-out" cross validation. That is, we regarded one participant as an end user and remaining participants as crowd users, and we computed the activity recognition performance of the end user's sensor data (test data). We evaluated the performance of our method by using error rate and error reduction rate. Error rate E_r is shown as formula (4) and error reduction rate E_{rr} is shown as formula (5).

$$E_r = 1 - Fmeasure \quad (4)$$

$$E_{rr} = 1 - \frac{E_r \text{ of proposed method}}{E_r \text{ of naive method}} \quad (5)$$

In this experiment, a naive method trains an activity recognition model (C4.5 decision tree) simply by using sensor data obtained from the crowd user as feature vectors.

4.3 Results

Error rates for each method obtained in the evaluation experiment are shown in Fig. 4. From the results of Fig. 4, our proposed method achieved 5.1% of error rate reduction in comparison with naive method.

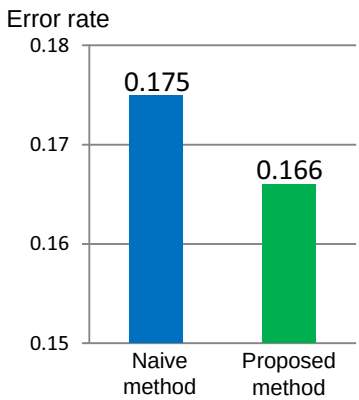


Figure 4. Error rates of naive method and proposed method

4.4 Discussion

From the results, we can say that we could improve the activity recognition accuracy by using the proposed method. Here we investigate the results for experiment participants with different physical characteristics in detail. Fig. 5 shows the average error reduction rate related to the participants with specific physical characteristics. The left graph shows the results related to writing dominant, whereas the right graph shows the results related to the height.

We firstly focus on writing dominant. The error reduction rate of right-handed participants (n=54) was 4.5%, and the one of left-handed participants (n=7) was 8.6%. In other words, impacts of our method on the left-handed participants was 1.9 times higher than that on the right-handed participants. We then focus on the height results. The error reduction rate of participants with less than 160cm tall (n=12) was 6.9%, the one of participants with 160-170cm tall (n=20) is 4.1%, and the one of participants with more than 170cm tall (n=29) was 5.1%. The error reduction rate of participants with less than 160cm tall (n=12) was 1.7 times higher than that of participants with 160-170cm tall (n=20). We can say that the proposed method is more effective for an end user with minor physical characteristics. This may be because most training data points, which are obtained from right-handed participants, had negative effects on test data from left-handed participants, and our method could filter out such unnecessary data points for left-handed participants.

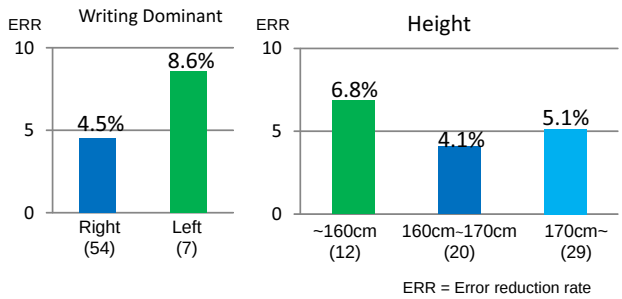


Figure 5. Error reduction rate of the participant with different physical characteristics

To investigate the effectiveness of our method (Fig. 4), we performed one-tailed t-test. That is, we compare the difference between the error rate of the naive and that of the proposed method. The results are shown in Table 3. Note that $Average$ is the average of error rate, $Variance$ is the variance of error rate, N is the number of samples (participants), t is t statistics, and P is the p-value of the t-test. As shown in Table 3, we confirmed a significant difference between the error rate of the naive and that of the proposed method by t-test ($p \leq .01$). In addition, to investigate the difference between the effect

of our method on right-handed participants and left-handed participants (right portion of Fig. 5), we performed one-tailed t-test. That is, we compare the difference between the error rate for the right-handed participants and that for the left-handed participants. The results is shown in Table 4. As shown in Table 4, we could not confirm a significant difference between the error rate for the right-handed participants and that for the left-handed participants ($p \geq .01$). This may be because our data set includes small numbers of left-handed participants (seven participants).

Table 3. T-test result related to naive method and proposed method

	Naive method	Proposed method
Average	0.174973	0.165985
Variance	0.00237	0.001997
N	61	61
t	3.812338	
P	0.000164	

Table 4. T-test result related to user groups with different writing dominant

Writing Dominant	Right	Left
Average	0.039042	0.083226
Variance	0.009964	0.00657
N	54	7
t	-1.12147	
P	0.133315	

5 Conclusion

This paper proposed an adaptive training data selection method with physical characteristics information for acceleration-based daily activity recognition. training an activity model using training data collected by the crowd reduces activity recognition accuracy for an end user's sensor data because the sensor data from other users who have different physical characteristics from the end user are also different from those from the end user. In the proposed method, by focusing on an area where data points belonging to different activity classes are mixed, we determine which data point is used to train the end user's activity recognition model using physical characteristics information. From the results, we could verify that our proposed method achieved higher activity recognition accuracy than naive method. We also found that the proposed method is more effective for an end user with minor physical characteristics.

Acknowledgment

The authors would like to thank Dr. Yasushi Sakurai of the NTT Communication Science Laboratories for providing acceleration data sets for evaluation experiment.

References

[1] T. Maekawa, Y. Yanagisawa, Y. Kishino, K. Ishiguro, K. Kamei, Y. Sakurai, and T. Okadome, "Object-based activity recognition with heterogeneous sensors on wrist," In Pervasive 2010, pages 246-264, 2010.

[2] R. Ohmura, N. Hashida, and M. Imai, " Preliminary evaluation of personal adaption techniques in accelerometer-based activity recognition," In Int'l Symp. on Wearable Computers: Late Breaking Results, 2009.

[3] K. Forster, A. Biasiucci, R. Chavarriage, J. del, R. Millan, D. Roggen, and G. Troster, "On the use of brain decoded signals for online user adaptive gesture recognition systems," In Pervasive 2010, pages 427-444, 2010.

[4] M. Stikic, K. Van Laerhoven, and B. Schiele, "Exploring semi-supervised and active leaning for activity recognition," In Int'l Symp. on Wearable Computers, pages 81-88, 2008.

[5] T. Huynh and B. Schiele, "Towards less supervision in activity recognition from wearable sensors," In Int'l Symp. on Wearable Computers, pages 3-10, 2006.

[6] G. Krassnig, D. Tantiger, C. Hofmann, T. Wittenberg, and M. Struck, "User-friendly system for recognition of activities with an accelerometer," In PervasiveHealth 2010, pages 1-8, 2010.

[7] L. Bao, and S. Intille, "Activity recognition from user-annotated acceleration data," In Proc. of Int'l Conf. on Pervasive Computing (PERVASIVE), pp.1-17, (2004).

[8] N. Japkowicz, et al., "Learning from imbalanced data sets: a comparison of various strategies," In Proc. of AAAI Workshop : Learning from Imbalanced Data Sets, pp.10-15 (2000).

[9] J.R. Quinlan, "C4. 5: programs for machine learning," Morgan Kaufmann, 1993.

Dynamic State Estimation for Vehicle Platoon System Based on Feedback Estimators

Hironori Suzuki† **Takashi Nakatsuji‡**

†School of Engineering, Nippon Institute of Technology, Japan
‡School of Engineering, Hokkaido University, Japan
viola@nit.ac.jp, naka@eng.hokudai.ac.jp

Abstract

A dynamic feedback system is developed for estimating the headway and velocity in a longitudinal three-vehicle platoon. The estimation system is modeled using a particle filter (PF) and an unscented Kalman filter (UKF) that estimate them by measuring the acceleration rate and/or velocity of probe vehicle(s) in the platoon. State equations are defined as a discrete conservation equation of headway and velocity, whereas the measurement equation is based on a conventional car-following model. The UKF and PF have the advantage of avoiding first-order approximation when implementing a filtering process to increase the estimation accuracy. Numerical analyses using real car-following data showed that the PF and UKF reduce the estimation errors in most cases at satisfactorily levels around 1m for headway and 1m/s for velocity estimations.

Keywords: Vehicle Platoon, State Estimation, Particle Filter, Unscented Kalman Filter, Feedback Estimation

1 Introduction

A vehicle platoon is a system that consists of multiple car-following vehicles moving longitudinally on an arterial road or freeway corridor. In order to achieve a high level of safety of vehicle platoon, not only the collision avoidance system of a single vehicle but also the cooperative control of multiple vehicles are required. Some studies have been proposed to evaluate and mitigate the risk of rear-end collision over a three-vehicle platoon by designing an appropriate human machine interface (HMI) [1, 2]. However, applying the HMI to a large platoon is still a difficult task because the sensing data are not obtained and shared in real-time by all platoon vehicles in the present car-following situation. Unfortunately, all the vehicles are not always quipped with a data collection system that can measure the headway and velocity. To bring a risk mitigation system to the real world in near future, an intelligent system is required that estimate the headway and velocity of all platoon vehicles indirectly from the measurement variables of some part of prove vehicles out of the platoon.

The authors have been attempting to develop a system of estimating the headway and velocity of a three-vehicle platoon using extended Kalman filter (EKF), a Neural Kalman Filter (NKF), Particle Filter (PF) and an Unscented Kalman Filter (UKF) [3, 4]. The study by Suzuki and Nakatsuji [5] showed that PF and UKF which require no first-order approximation in the model development performed the best among the other estimators such as EKF and NKF. Especially, PF is well known as a powerful tool and has been widely applied to various dynamic estimations such as traffic states estimation [5], estimating visual shape and motion [6], mobile robot control [7], target tracking [8] and so on.

As demonstrated above, PF performs very well in various non-linear estimation problems including the identification of headway and velocity of vehicle platoon. However, our previous study of vehicle platoon estimation confirmed the performance of PF in the simulated artificial car-following data only. Applicability of PF still remains unknown for real car-following situations.

This research aims to evaluate whether the PF and UKF estimate the velocity and headway of a three-vehicle platoon accurately even in real car-following scenarios. Also, performance of these two estimators is evaluated and compared by changing the number of measurement variables. Here, in this paper, only PF and UKF are selected as the feedback estimators since it has been proved that they performed the best compared to the existing approaches such as EKF and NKF.

2 Theoretical Background

2.1 State space model

The following state and measurement equations are defined as the state-space model, which is required for feedback estimation:

$$\mathbf{x}_k = \mathbf{f}\left(\mathbf{x}_{k-1}\right) + \mathbf{v}_{k-1}, \quad (1)$$

$$\mathbf{y}_k = \mathbf{g}\left(\mathbf{x}_k\right) + \mathbf{w}_k, \quad (2)$$

where

- $\mathbf{x}_k$: state variables at time k ,
- $\mathbf{y}_k$: measurement variables,
- $\mathbf{v}_k$: system error,
- $\mathbf{w}_k$: measurement error,
- $\mathbf{f}$, $\mathbf{g}$: possible non-linear functions.

Let $\mathbf{y}_{1:k}$ and $\mathbf{x}_{1:k}$ denote as a set of all available measurement and state variables up to time k given by:

$$\mathbf{y}_{1:k} \equiv \left\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\right\}, \quad (3)$$

and

$$\mathbf{x}_{1:k} \equiv \left\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\right\}. \quad (4)$$

The estimation problem is to calculate the posterior probability $p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k}\right)$ when giving a set of all measurement $\mathbf{y}_{1:k}$. If measurement $\mathbf{y}_k$ is available, the posterior is updated through the Baye's rule:

$$p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k}\right) = \frac{p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k-1}\right) \cdot p\left(\mathbf{y}_k \middle| \mathbf{x}_k\right)}{p\left(\mathbf{y}_k \middle| \mathbf{y}_{1,k-1}\right)}. \quad (5)$$

Here, $p\left(\mathbf{y}_k \middle| \mathbf{x}_k\right)$ is a likelihood function which describes the likelihood of $\mathbf{x}_k$ when giving the measurement $\mathbf{y}_k$. $p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k-1}\right)$ is a prior probability of $\mathbf{x}_k$ given by:

$$p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k-1}\right) = \int p\left(\mathbf{x}_k \middle| \mathbf{x}_{k-1}\right) \cdot p\left(\mathbf{x}_{k-1} \middle| \mathbf{y}_{1,k-1}\right) d\mathbf{x}_{k-1}. \quad (6)$$

Also, $p\left(\mathbf{y}_k \middle| \mathbf{y}_{1,k}\right)$ is defined as:

$$p\left(\mathbf{y}_k \middle| \mathbf{y}_{1,k-1}\right) = \int p\left(\mathbf{y}_k \middle| \mathbf{x}_k\right) \cdot p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k-1}\right) d\mathbf{x}_k. \quad (7)$$

$p\left(\mathbf{x}_k \middle| \mathbf{x}_{k-1}\right)$ in (4) and $p\left(\mathbf{y}_k \middle| \mathbf{x}_k\right)$ in (5) can be given by the state and measurement equations (1) and (2).

There is a strong restriction that the posterior and prior probability density functions (PDF) $p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k}\right)$ and $p\left(\mathbf{x}_k \middle| \mathbf{y}_{1,k-1}\right)$ should be Gaussian in a conventional Kalman filter, whereas the Particle Filter (PF) does not require any assumptions or analytical functions for the PDFs.

2.2 Particle filter (PF) [9, 10, 11]

The key idea of particle filter (PF) is to represent the required posterior PDF by a set of random samples with associated weights and to compute the estimates based on these samples and weights [10].

Let $\left\{\left(x_{1:k-1}^{(i)}, w_{1:k-1}^{(i)}\right)\right\}_{i=1}^{M}$ denote as a set of particles and its associated weights. It is assumed that $\left\{\left(x_{1:k-1}^{(i)}, w_{1:k-1}^{(i)}\right)\right\}_{i=1}^{M}$ is given from the posterior PDF at time $k-1$. When giving the measurement $\mathbf{y}_k$ at time step k , the PF is to compute and update the particles and weights from $\left\{\left(x_{1:k-1}^{(i)}, w_{1:k-1}^{(i)}\right)\right\}_{i=1}^{M}$ to $\left\{\left(x_{1:k}^{(i)}, w_{1:k}^{(i)}\right)\right\}_{i=1}^{M}$.

The weight is updated through the appropriate process called Sequential Importance Sampling (SIS). The SIS generates the particles based on the proposal distribution $q\left(\mathbf{x}\right)$ which is different from the objective distribution $p\left(\mathbf{x}\right)$ and then gives each particle the appropriate weight so as to make the particles to be close to the objective PDF.

The update process of weight is given by:

$$\tilde{w}_k^{(i)} = w_k^{(i)} \frac{p\left(\tilde{\mathbf{x}}_k^{(i)} \middle| \mathbf{x}_{k-1}^{(i)}\right) p\left(\mathbf{y}_k \middle| \mathbf{x}_k^{(i)}\right)}{q\left(\tilde{\mathbf{x}}_k^{(i)} \middle| \mathbf{x}_{1:k-1}^{(i)}, \mathbf{y}_{1,k}\right)}. \quad (8)$$

$$\left(i = 1, 2, \ldots, M\right)$$

The PF is the feedback estimator that theoretically places the random particles in the probability field to yield the accurate posterior PDF based on the Baye's theory. It is guaranteed that the estimates through the filtering process are suboptimal. The estimation process of PF is as follows:

1. Initial sampling: Generate the initial particles $\hat{\mathbf{x}}_0^{(i)}$ $\left(i = 1, 2, \ldots, M\right)$ at $k = 0$ from the initial

probability density function $p(\hat{\mathbf{x}}_0)$.

2. Importance sampling: Generate one-step prediction of samples $\tilde{\mathbf{x}}_k^{(i)}$ and $\tilde{\mathbf{y}}_k^{(i)}$ based on the state and measurement equations:

$$\tilde{\mathbf{x}}_k^{(i)} = \mathbf{f}\left(\hat{\mathbf{x}}_{k-1}^{(i)}\right) + \mathbf{v}_{k-1}^{(i)}, \tag{9}$$

$$\tilde{\mathbf{y}}_k^{(i)} = \mathbf{g}\left(\tilde{\mathbf{x}}_i^{(i)}\right) + \mathbf{w}_k. \tag{10}$$

3. Importance weight: Calculate the importance weight $\tilde{w}_k^{(i)}$ for each sample by the following equation:

$$\tilde{w}_k^{(i)} = w_k^{(i)} \frac{p\left(\tilde{\mathbf{x}}_k^{(i)} \middle| \mathbf{x}_{k-1}^{(i)}\right) p\left(\mathbf{y}_k \middle| \tilde{\mathbf{x}}_k^{(i)}\right)}{q\left(\tilde{\mathbf{x}}_k^{(i)} \middle| \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1,k}\right)}. \tag{11}$$

$$(i = 1, 2, ..., M)$$

Larger weight is given to particles for which the prediction $\tilde{\mathbf{y}}_k^{(i)}$ is closer to the actual measurement $\mathbf{y}_k$.

4. Resampling: The particle $\tilde{\mathbf{x}}_k^{(i)}$ is updated into $\hat{\mathbf{x}}_0^{(i)}$ in proportion to the importance weight $\hat{w}_k^{(i)}$.

5. Normalize weight: Normalize the importance weight to $\hat{w}_k^{(i)}$ by

$$\hat{w}_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=1}^{M} \tilde{w}_k^{(j)}}. \tag{12}$$

6. Estimation: From the updated particles $\hat{\mathbf{x}}_0^{(i)}$, compute the estimate by

$$\hat{\mathbf{x}}_k \approx \frac{1}{M} \sum_{i=1}^{M} \hat{\mathbf{x}}_k^{(i)} \left(= \sum_{i=1}^{M} \hat{w}_k^{(i)} \tilde{\mathbf{x}}_k^{(i)} \right). \tag{13}$$

2.3 Unscented Kalman filter (UKF)

The UKF is another of the family of *derivative-free* Kalman filters, which require no calculation of partial derivative of state and measurement equations [12]. The conventional extended Kalman filter (EKF) provides the first-order approximation to the optimal estimates, whereas the UKF captures the posterior mean and covariance accurately to their second-order Taylor expansions using a minimal set of carefully chosen sample points called sigma points [12].

The first step to apply the UKF is to generate the sigma point as:

$$\boldsymbol{\sigma}_{i,P} = \left(\sqrt{(N+\lambda)\mathbf{P}_k}\right)_i, \tag{14}$$

$$\hat{\Phi}_{0,k-1} = \hat{\mathbf{x}}_0, \tag{15}$$

$$\hat{\Phi}_{i,k-1} = \hat{\mathbf{x}}_{k-1} + \boldsymbol{\sigma}_{i,P} \quad (i = 1, \ldots, N), \tag{16}$$

$$\hat{\Phi}_{i,k-1} = \hat{\mathbf{x}}_{k-1} - \boldsymbol{\sigma}_{i-N,P} \quad (i = N+1, \ldots, 2N), \tag{17}$$

$$\tilde{\boldsymbol{\Phi}}_k = \mathbf{f}\left(\hat{\boldsymbol{\Phi}}_{k-1}\right), \tag{18}$$

where

$\boldsymbol{\sigma}_i$: i-th column vector of the lower triangle matrix when applying the Cholesky decomposition to $\mathbf{P}_k$,

N: number of state variables,

λ: scaling parameter.

One-step prediction of state variables and the error covariance are then given by:

$$\tilde{\mathbf{x}}_k = \sum_{i=0}^{2N} h_i \tilde{\Phi}_{i,k}, \tag{19}$$

$$\mathbf{M}_k^{xx} = \sum_{i=0}^{2N} h_i \left[\tilde{\Phi}_{i,k} - \tilde{\mathbf{x}}_k\right]\left[\tilde{\Phi}_{i,k} - \tilde{\mathbf{x}}_k\right]^T + \mathbf{V}_{k-1}, \tag{20}$$

where h_i is the weight defined by

$$h_i = \frac{1}{2(N+\lambda)}. \tag{21}$$

The same procedure is applied to compute $\tilde{\mathbf{y}}_k$, $\mathbf{M}_k^{yy}$, and $\mathbf{M}_k^{xy}$ by defining the sigma vectors $\tilde{\boldsymbol{\Psi}}_k$ and $\tilde{\boldsymbol{\Omega}}_k$ as follows:

$$\boldsymbol{\sigma}_{i,M} = \left(\sqrt{(N+\lambda)\mathbf{M}_k}\right)_i, \tag{22}$$

$$\tilde{\Omega}_{0,k-1} = \tilde{\mathbf{x}}_0, \tag{23}$$

$$\tilde{\Omega}_{i,k} = \tilde{\mathbf{x}}_k + \boldsymbol{\sigma}_{i,M} \quad (i = 1, \ldots, N), \tag{24}$$

$$\tilde{\Omega}_{i,k} = \tilde{\mathbf{x}}_k - \boldsymbol{\sigma}_{i,M} \quad (i = N+1, \ldots, 2N), \tag{25}$$

$$\tilde{\boldsymbol{\Psi}}_k = \mathbf{g}\left(\tilde{\boldsymbol{\Omega}}_k\right), \tag{26}$$

$$\tilde{\mathbf{y}}_k = \sum_{i=0}^{2N} h_i \tilde{\Psi}_k, \tag{27}$$

$$\mathbf{M}_k^{yy} = \sum_{i=0}^{2N} h_i \left[\tilde{\Psi}_{i,k} - \tilde{\mathbf{y}}_k\right]\left[\tilde{\Psi}_{i,k} - \tilde{\mathbf{y}}_k\right]^T + \mathbf{W}_{k-1}, \tag{28}$$

$$\mathbf{M}_k^{xy} = \sum_{i=0}^{2N} h_i \left[\tilde{\Phi}_{i,k} - \tilde{\mathbf{x}}_k\right]\left[\tilde{\Psi}_{i,k} - \tilde{\mathbf{y}}_k\right]^T. \tag{29}$$

The Kalman gain is computed by (30) from the error covariance $\mathbf{M}_k^{yy}$ and $\mathbf{M}_k^{xy}$. Also, the optimal estimate $\hat{\mathbf{x}}_k$ and $\mathbf{P}_k$ are updated through (31) and (32).

$$\mathbf{K}_k = \mathbf{M}_k^{xy}\left(\mathbf{M}_k^{yy}\right)^{-1}, \tag{30}$$

$$\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k + \mathbf{K}_k\left(\mathbf{y}_k - \tilde{\mathbf{y}}_k\right). \tag{31}$$

$$\mathbf{P}_k = \mathbf{M}_k^{xx} - \mathbf{K}_k\mathbf{M}_k^{xx}\mathbf{K}_k^T. \tag{32}$$

3. State Space Model for Vehicle Platoon

Assuming that three vehicles form a longitudinal platoon system, as depicted in Figure 1, the dynamics of headway and velocity of each platooned vehicle are defined as

$$\ell_k^i = \ell_{k-1}^i + \left[v_{k-1}^{i-1} - v_{k-1}^i\right]\Delta t, \tag{33}$$

and

$$v_k^i = v_{k-1}^i + a_{k-1}^i \Delta t, \tag{34}$$

where

ℓ_k^i : headway of vehicle i at time k,
v_k^i : velocity,
a_k^i : acceleration rate,
Δt : discrete time step.

Figure 1. Three-vehicle platoon

The acceleration rate in Equation (34) is given by a conventional car-following model, which is well known as the Gazis-Herman-Rotery (GHR) model [13]:

$$a_{k+T}^i = \alpha \frac{\left(v_k^i\right)^n}{\left(\ell_k^i\right)^m}\left[v_k^{i-1} - v_k^i\right], \tag{35}$$

where

α, m, n : model parameters,
T : vehicle reaction time.

For simplicity, the reaction time is assumed to be zero, i.e. $T = 0$, under the assumption that errors due neglecting T will be minimized by the feedback process of the estimator. For the same reason, it is assumed that $\alpha = m = n = 1$. Substituting Equation (35) into (34) redefines the velocity as

$$v_k^i = v_{k-1}^i + \alpha \frac{\left(v_{k-1}^i\right)^n}{\left(\ell_{k-1}^i\right)^m}\left[v_{k-1}^{i-1} - v_{k-1}^i\right]\Delta t. \tag{36}$$

Equations (33) and (36) are regarded as the state equations, whereas Equation (35) can be used as the measurement equation for the state space model of a vehicle platoon. If velocity is also chosen as a measurement variable, the additional measurement equation is:

$$v_k^i = v_k^i. \tag{37}$$

When assuming that the 3rd vehicle is the only probe car, i.e. the only car equipped with a sensing system, the estimation problem is reduced to precisely estimating the state variables such as headway and velocity of the 2nd and 3rd vehicles by observing the measurement variables of acceleration and/or velocity of the 3rd vehicle. When observing acceleration only, the state and measurement variables are:

$$\mathbf{x}_k = \left[\ell_2, v_2, \ell_3, v_3\right]_{(k)}^T, \mathbf{y}_k = \left[a_3\right]_{(k)}^T, \tag{38}$$

otherwise,

$$\mathbf{x}_k = \left[\ell_2, v_2, \ell_3, v_3\right]_{(k)}^T, \mathbf{y}_k = \left[v_3, a_3\right]_{(k)}^T. \tag{39}$$

It is preferable that the velocity of the 1st vehicle is also added to the measurement variables. However, v_k^1 is not included in the measurement variables in this analysis, but explicitly given to the platoon system since the transition of v_k^1 is not described and modeled in the state equation.

4. Numerical Analyses

4.1. Preparing Real Car-Following Data Sets

Real car-following data including headway ℓ_k^i, velocity v_k^i and acceleration rate a_k^i of three vehicles were collected in a field test using a test truck of the Japan Automobile Research Institute. In the test, a three-vehicle platoon travelling at a steady speed of around 60 km/h was made to decelerate and come to a complete stop, and this process was repeated. The deceleration rate was random from 1 to 5 m/s^2.

Thirteen scenarios (S1 to S13) which are suitable for the analysis were selected and used for the evaluation.

4.2 Evaluation Result for Measuring Acceleration Only

Figures 2 and 3 compare the headway and velocity estimates of scenario S2 among two estimators, PF,UKF and no-filter case. PF and UKF yield more accurate estimates than no-filter case for all state variables. Especially in the headway estimation of 2nd and 3rd vehicles, PF and UKF reduced the unexpected under estimation that was seen in no-filter case. Both PF and UKF are very close to the observed target, but UKF is more stable and accurate than PF. UKF seems to absorb the system and measurement noises to yield stable estimates.

Figures 4 and 5 depict the mean and standard deviation of root mean square errors (RMSE) of four state variables among thirteen scenarios. The t-test showed that mean errors of both headway estimates are statistically smaller than the no-filter case at 1% confidence level. The absolute error around 1 m, which is equivalent to the error by laser radar, is also acceptable as the satisfactorily level.

In the velocity estimations, however, no statistical difference is observed except the PF in the velocity estimates of 2nd vehicle although PF and UKF seem to decrease the mean errors. But, the absolute mean error around 0.8 to 1.2 m/s can be considered as the acceptable level for the collision risk evaluation.

Measuring the acceleration rate of 3rd vehicle is enough to yield accurate estimates of both headway and velocity. This is because the platoon size is very small and the three vehicles are travelling in an appropriate car-following situation without any unexpected events.

4.3 Evaluation Results for Measuring both Acceleration and Velocity

The Estimates by PF and UKF when measuring both acceleration and velocity are depicted in Figures 6 and 7 for the same scenario S2. This scenario is a nominal case that additional measurement variable increases the estimation accuracy. Unfortunately, however, there are some scenarios where the estimates are slightly worse by adding the velocity of 3rd vehicle as a measurement variable.

Figures 8 and 9 evaluate the estimation errors in average. The mean errors still remain small at the satisfactorily levels for both headway and velocity estimations. When comparing Figures 4 and 8 or 5 and 9, however, there are some cases where the error is increased by adding the velocity of 3rd

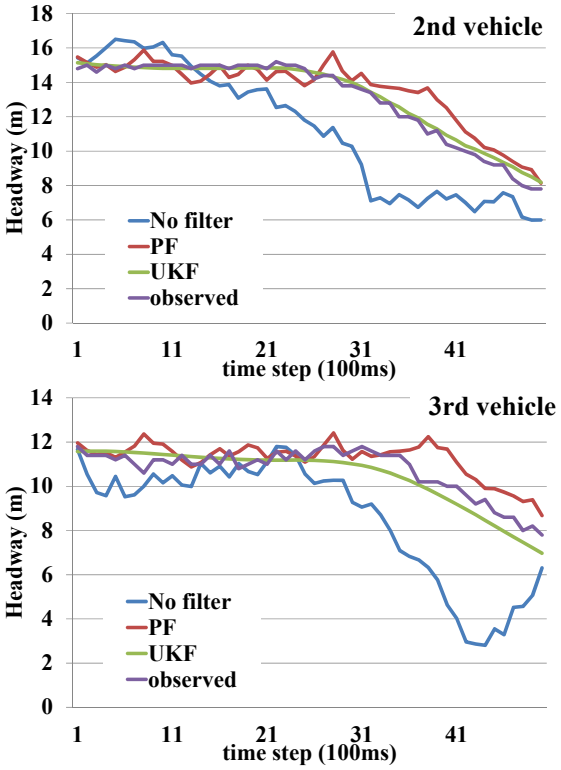


Figure 2. Headway estimates when observing acceleration only (scenario S2, upper: 2nd vehicle, lower: 3rd vehicle)

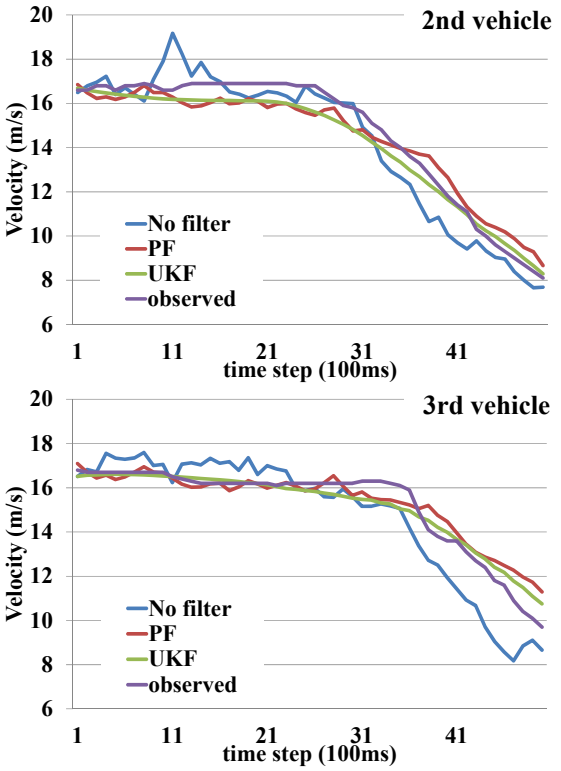


Figure 3. Velocity estimates when observing acceleration only (scenario S2, upper: 2nd vehicle, lower: 3rd vehicle)

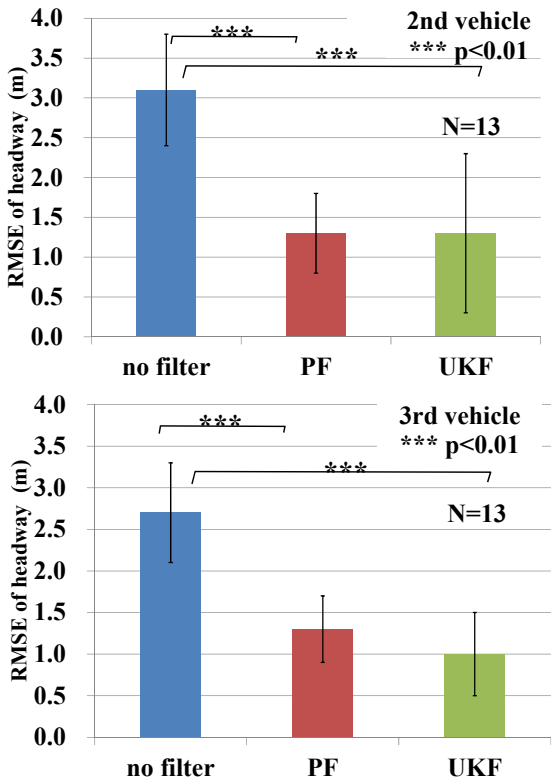


Figure 4. RMSE of headway estimates when observing acceleration only (upper: 2nd vehicle, lower: 3rd vehicle)

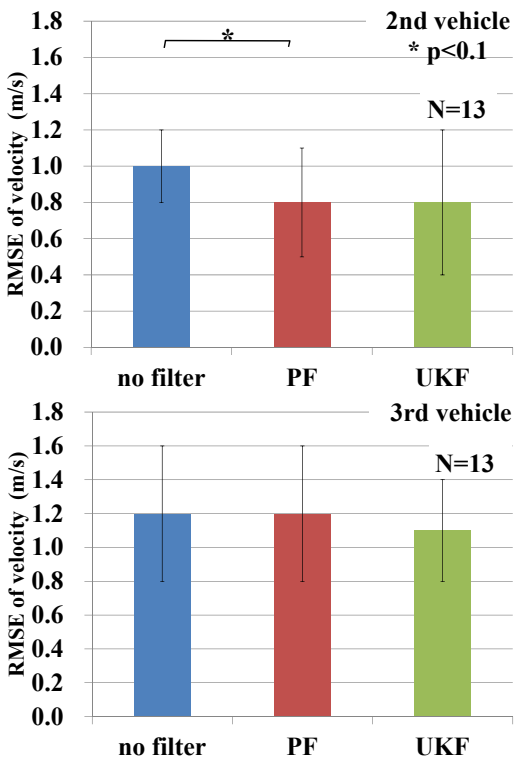


Figure 5. RMSE of velocity estimates when observing acceleration only (upper: 2nd vehicle, lower: 3rd vehicle)

vehicle as a measurement variable. The reason why the accuracy is worse when adding the measurement variable is not identified yet. The appropriate measurement variables that give us optimum estimates are depending on the platoon size or the position of probe cars. We need more sophisticated investigation to find out them by increasing the platoon size or changing the location of probe cars. If limited in this analysis, however, measuring the acceleration rate of the 3rd vehicle is enough to estimate the headway and velocity accurately.

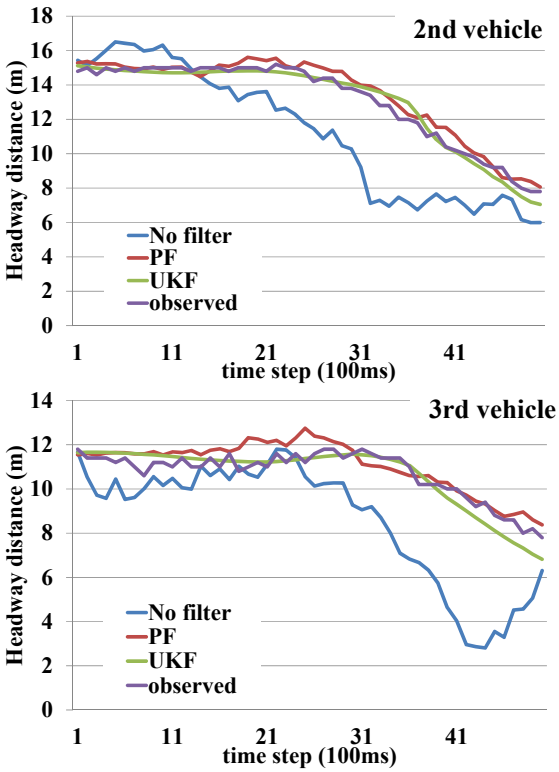


Figure 6. Headway estimates when observing acceleration and velocity of 3rd vehicle (scenario S2; upper: 2nd vehicle, lower: 3rd vehicle)

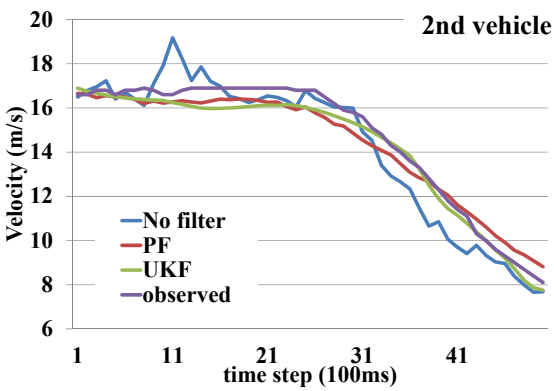


Figure 7. Velocity estimates when observing acceleration and velocity of 3rd vehicle (scenario S2)

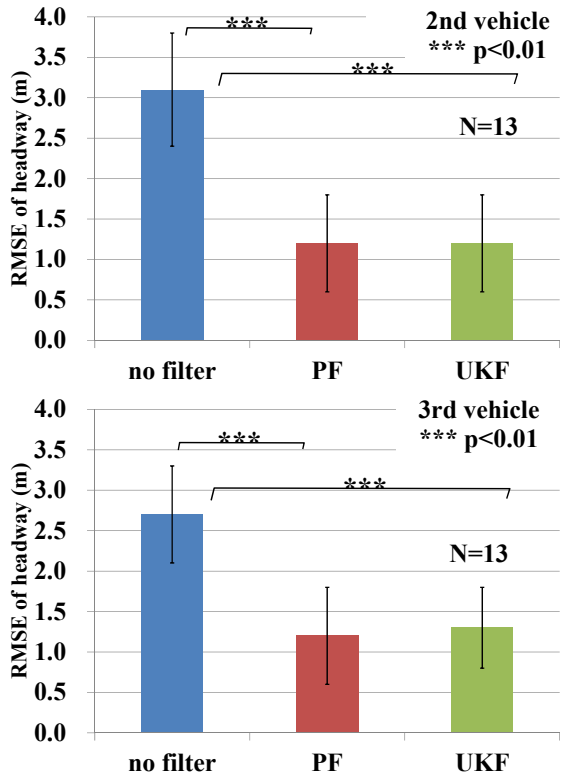


Figure 8. RMSE of headway estimates when observing acceleration and velocity of 3rd vehicle (upper: 2nd vehicle, lower: 3rd vehicle)

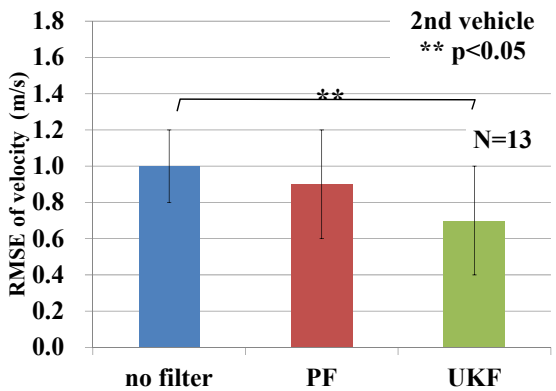


Figure 9. RMSE of velocity estimates when observing acceleration and velocity of 3rd vehicle

Concluding Remarks

This research attempts to develop an intelligent feedback system for estimating dynamic states of three-vehicle platoon. Particle filter (PF) and unscented Kalman filter (UKF) are applied to estimate the headway and velocity indirectly from the acceleration rate and/or velocity of a prove vehicle. Numerical analyses showed that even when using real car-following data, the estimation accuracy is as low as a satisfactorily level compared to the case where no filtering process is applied. Also, it was found that only the acceleration rate of prove car is enough to provide an appropriate estimation precision for the headway and velocity estimations. However, this finding is limited only in the smallest platoon which consists of solely three vehicles.

More number of measurement variables or prove cars may be required when the platoon size is increased. Further work should be carried out to apply the estimation to a larger platoon system with longer time period to evaluate whether the proposed algorithm by PF or UKF is applicable to the real world car-followings.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number C24510231. The authors wish their sincere appreciation to Japan Automobile Research Institute (JARI) for giving us a data set of real car-following data.

References

[1] S. Takada, T. Hiraoka, H. Kawakami. Effect of forward obstacles collision warning system based on deceleration for collision avoidance on driving behavior, *Proceedings of 19th World Congress on Intelligent Transport Systems*, CD-ROM, 2012.

[2] Y. Marumo, K. Tanaka, H. Suzuki. Assistance system to predict driving behavior considering information on pre-preceding vehicle, *Second International Symposium on Future Active Safety Technology*, *FAST-zero*, 2013. (in printing)

[3] H. Suzuki, T. Nakatsuji. Neural Kalman filter for estimating dynamic velocity and headway distance in vehicle platoon system -validity of neural Kalman filter under unexpected car-following situation-", *Proceedings of IEEE SMC 2012*, 2012.

[4] H. Suzuki and T. Nakatsuji. Feedback estimators for identifying headway distance and velocity of longitudinal platooned vehicles, *Journal of Eastern Asia Society for Transport Studies (EASTS)*, 2013. (in printing)

[5] T. Kobayashi and T. Nakatsuji. Dual estimators for dynamic traffic states estimations and parameters, *Proceedings of Infrastructure Planning and Management, Japanese Society of Civil Engineers*, Vol. 46, 2012. (in Japanese)

[6] A. Blake, M. Isard, J.MacCormic, Statistical models of visual shape and motion, *in Sequential Monte Carlo Methods in Practice*, 339-358,

Springer, 2001.
[7] D. Fox, S. Thrun, W. Burgard, F. Dellaert, Particle filters for mobile robot localization, *in Sequential Monte Carlo Methods in Practice*, 401-428, Springer, 2001.
[8] S. McGinnity, G.W. Irwin, Maneuvering Target tracking using a multiple-model bootstrap filter, *in Sequential Monte Carlo Methods in Practice*, 479-498, Springer, 2001.
[9] A. Doucet , A M. Johansen. A tutorial on particle filtering and smoothing: fifteen years later, *Oxford Handbook of Nonlinear Filtering*, 2011.
[10] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp., A tutorial on particle filters for online nonlinear/non-gGaussian Bayesian tracking, *IEEE Transcactions on Signal Processing*, 50(2): 174-188, 2002.
[11] N. Ikoma. Sequential Monte Carlo method and particle filter, Chapter 11, Vol. III, *Statistical Science of 21 Century, HP edition*, Japan Statistical Society, 2012. (in Japanese)
[12] S. Haykin, S. Kalman filtering and neural networks, John Wiley & Sons, 2001.
[13] D. L. Gerlough, M. J. Huber. Traffic flow theory–a monograph, *Special Report 165, Transportation Research Board*, 1975.

New DEA Framework: Evaluating I/O Efficiency Structures

Kazushige Inoue† **Kenji Fukuoka†** **Shingo Aoki‡** **Hiroshi Tsuji†**

†Osaka Prefecture University, Osaka, Japan
‡Hiroshima Institute of Technology, Hiroshima, Japan
{inoue, fukuoka}@mis.cs.osakafu-u.ac.jp
s.aoki.sm@it-hiroshima.ac.jp
tsuji@osakafu-u.ac.jp

Abstract

We propose a framework for analyzing structures of efficiency. The proposed I/O efficiency structure helps analysts evaluate items visually what I/O items are characteristic of the DMUs and their level of efficiency. The analysis consists of four steps. In step 1, combinations of I/O items are regarded as nodes, and a hierarchical structure is constructed on the basis of nodes' inclusive relations. In step 2, efficiency values are calculated for each node. In step 3, nodes are integrated on the basis of efficiency values. In step 4, the integrated nodes are plotted in the order of their efficiency values. Applying the proposed framework, analysts determine subsequent efficient items and key items that improve efficiency values when combined with other items. We tested the proposed framework by applying it to a sample dataset with two input items, four output items, and eighteen DMUs.

Keywords: DEA, Linear Programming, Decision Support, I/O Efficiency Structure

1 Introduction

The continued development of information technology provides business managers with copious amounts of data, which they can use to make decisions. However, it is difficult for business managers to make decisions using the original data; therefore many methods such as operations research and data mining have been developed that help extract information from the raw data. For example, Data Envelopment Analysis (DEA) is a useful tool [1] that has been applied in various fields [2][9], and many models based on DEA have been developed [8][10][11][12].

In DEA, companies are regarded as Decision Making Units (DMUs) that produce multiple outputs from multiple inputs, and the input/output (I/O) ratio is calculated as an efficiency value. DMUs are analyzed by an evaluation axis that considers features of DMUs and vary according to them. Therefore, DEA helps analysts evaluate DMUs relatively with multiple input and output items, and it has been applied to many fields – for example, evaluate banks, stores, and universities [3][5][6][7]. Furthermore, DEA has given analysts the chance to identify the strengths of DMUs using variables calculated in DEA [2].

Moreover, to evaluate DMUs from various viewpoints, a hierarchical approach was proposed [3][4]. In this approach, analysts constructed a hierarchical structure by resolving evaluation perspectives. Each resolved perspective was viewed as a node and assigned I/O items to evaluate DMUs from the each perspectives. DMUs were analyzed from multiple-viewpoints based on the efficiency values calculated for each perspective. However, because the conventional approach evaluates DMUs with a hierarchical structure constructed on the basis of analysts' intention and combinations of I/O items that influence the efficiency values vary for the DMUs, structures of efficiency could not be evaluated with this approach.

Thus, we define structures of efficiency as "I/O efficiency structures" and propose a framework to analyze them on the basis of the conventional hierarchical approach. The proposed approach comprises the following four steps.

- Step 1: All combinations of I/O items are regarded as nodes, and the hierarchical structure is constructed on the basis of inclusive relationship among these nodes.
- Step 2: Efficiency values are calculated for each node.
- Step 3: The nodes are integrated together.
- Step 4: The nodes are plotted based on the

basis of calculated efficiency values.

To examine the proposed framework, a numerical experiment with two input and four output items and eighteen DMUs is carried out. This paper introduces *t*he conventional hierarchical approach and its problem in section 2. Then, to overcome the problem, we present a framework in section 3. Section 4 shows the utility of the proposed framework through numerical experiments. Finally, we conclude our research in section 5.

2 Conventional Hierarchical Approach

2.1 Outline of the Conventional Hierarchical Approach

Applying DEA which is introduced in *Appendix*, analysts can evaluate the efficiency and its strengths with multiple I/O items. Moreover, a hierarchical approach was developed to analyze efficiency from multiple-viewpoints [3][4]. In this approach, analysts construct a hierarchical structure by resolving evaluation perspectives. To evaluate DMUs from the perspectives of the nodes, I/O items are assigned to each node on the basis of two rules. The first rule is that the highest node represents an evaluation based on all I/O items. The second rule is that if there are two nodes, they are connected; thus, items of one node will include those of the other. In addition, we call the former the "parent node" and the latter the "child node" and the more items the nodes have, the higher their layer. Efficiency values are calculated for each *perspective* of the *hierarchical structure*, and because the efficiency values of the original dataset are not less than those of the selected dataset, values of the parent nodes do not exceed those of the child nodes. Furthermore, nodes belonging to the higher layer are more efficient.

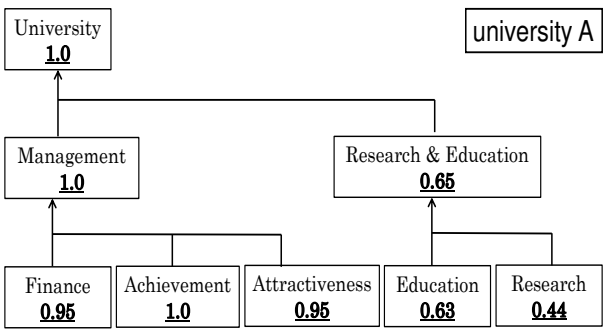


Figure. 1. Evaluation of University A by the

In [3], this approach was applied to evaluate universities in Japan. Figure 1 shows *the hierarchical structure* based on the evaluation *perspectives* and their efficiency values. Authors made a *hierarchical structure*, assigned I/O items to each node, and calculated the efficiency values. They determined that university A is efficient in "Management" – particularly with regards to "Achievement" – but it is inefficient in "Research & Education" – particularly with regards to "Research". While the efficiency value of "University" is calculated with all I/O items in the traditional DEA, the conventional hierarchical method allows analysts to evaluate DMUs from multiple-viewpoints.

2.2 The Problem of the Conventional Hierarchical Approach

As the structure of efficiency differs from that of DMUs and the conventional approach evaluates DMUs based on a hierarchical structure constructed by analysts' intention, it is difficult to analyze the structure of efficiency for each DMU. Therefore, the following two features cannot be found using the conventional approach.

a. Subsequent efficient combinations of I/O items: Assuming that there is a DMU whose output 1 is the most efficient and output 2 is the second most efficient, this DMU is evaluated as efficient by the conventional DEA because of only output 1 and the efficiency considering output 2 cannot be evaluated. Therefore, subsequent efficient combinations, such as combinations including output2, and their level of efficiency cannot be evaluated.
b. Key items for efficiency: Assuming that there is a DMU whose output 2 is inferior to other DMUs, there are some cases in which combining output 2 with other output items increases efficiency. However, the weights in DEA are assigned to items based on strength, and the item whose weight is low is evaluated as having very little strength. Therefore, the analysts cannot evaluate whether an item is key item that can improve efficiency values when combined with other items.

3 The Proposed Framework

3.1 Step 1: Constructing the Hierarchical Structure

In step 1, all combinations of I/O items are considered as nodes, and *the hierarchical structure* is constructed on the basis of following two rules: (1) Parent nodes include a child node as is the case in the conventional approach and (2) each node includes at least one input and one output item.

If the number of input items is m and the number of output items is s, then the number of nodes is calculated by $(2m - 1)(2s - 1)$. When a node whose number of input items is i ($i = 1,2,\ldots,m$) and whose number of output items is r ($r = 1,2,\ldots,s$) is expressed as node_{i+r}, it is linked to the node_{i+r+1} that satisfies $n(\text{node}_{i+r} \cap \text{node}_{i+r+1}) = i+r$. In this situation, node_{i+r} is a child node and the node_{i+r+1} is a parent node; a *hierarchical structure* is constructed on these two rules: (1) The highest layer, the $m+s$ th layer, has a node with all I/O items, (2) the lowest layer, the *1*st layer, has $m*s$ nodes with one input and one output item.

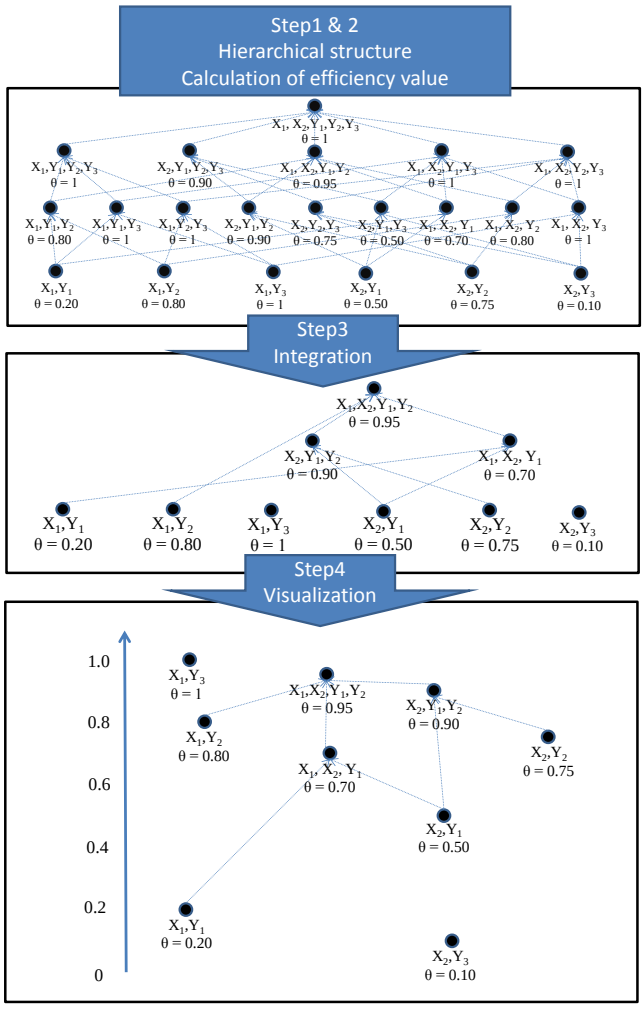


Figure. 2. An Image of Result Calculated by the Proposed Framework

3.2 Step 2: Calculating Efficiency Values

In step2, the efficiency value is calculated for each node. Because items of parent nodes are a combination of a child node and an added item, an efficiency value of the parent node is not less than that of the child node, as is the case with the conventional approach. If an efficiency value becomes higher than that of the child node, the combination of the child node and the added item becomes a characteristic of the DMU. However, if the efficiency value is equal to that of the child node, the combination is not a characteristic of the DMU. In particular, if there are some efficient nodes, all parent nodes linked to the efficient child nodes are then efficient, and added items cannot further influence the efficiency.

Figure 2 shows the proposed framework with two input and three output items. A circle represents the node and items included, and the efficiency value is below. In this example, the number of nodes is calculated as $(2^2-1)(2^3-1) = 21$ and the top figure shows the hierarchical structure. The lowest layer have six nodes with one input and one output item and the highest layer has one node with all I/O items. Efficiency values of parent nodes are not less than those of child nodes, and all parent nodes linked to an efficient node (X_1, Y_3) are efficient.

In conventional DEA, while the most efficient combination can be evaluated from weights, efficiency values of other combinations cannot be evaluated. In the proposed framework, (X_1, Y_3) is the most efficient combination, which makes this DMU efficient. Moreover, it can be determined that the efficiency value of (X_1, X_2, Y_1, Y_2), which does not include (X_1, Y_3), is 0.95. This combination is the second most efficient, which analysts could not find with the conventional approach. Furthermore, when (X_2, Y_1) is combined with X_1, the efficiency value changes from *0.50* to *0.70*, and when (X_2, Y_1) is combined with Y_2, the efficiency value changes to *0.90*. Therefore, although the efficiency value is low, efficiency values are improved if these items are combined with other items. However, as the number of I/O items increases, the number of

layers increases proportionally, and the number of nodes increases exponentially; thus, it can be difficult to analyze the I/O efficiency structure from hierarchical structure.

3.3 Step 3: Integrating Nodes

Focusing on the fact that when an efficiency value of the parent node is equal to that of the child node, the combination of parent node's items is not a characteristic for efficiency; thus, we integrate nodes based on efficiency values to decrease the number of nodes.

Then, if efficiency values of the parent nodes are equal to those of child nodes, we integrate parent nodes with the child node. We retain links that combine parent nodes and their parent nodes and clear links that combine parent nodes and their child nodes. However, if there are some child nodes whose efficiency values are equal to that of a parent node, we do not combine them.

On the basis of these rules, we integrate nodes from first layer to $m+s-2$ th layer. Because integrated nodes represent the characteristic combination of I/O items, analysts can visually evaluate what items are the characteristic of the DMUs for the efficiency values.

3.4 Step 4: Visualizing the I/O Efficiency Structure

While analysts can visually evaluate what items are the characteristic of the DMUs for efficiency values, they cannot visually evaluate the efficiency of the combination of items from the hierarchical structure. Thus, we plot integrated nodes with their links in order of efficiency values and visualize the I/O efficiency structure. Visualization helps analysts evaluate items that are characteristics of DMUs and determine their level of efficiency.

Let us explain the process of step 3 and step 4 shown in the middle and bottom figure of Fig. 2. The middle figure of Fig. 2 shows the *hierarchical structure* whose nodes are integrated. In this example, as the efficiency values of (X_1, Y_3) is *1*, nodes including (X_1, Y_3) are integrated to (X_1, Y_3). Because the efficiency value of (X_2, Y_1, Y_2, Y_3) equals that of (X_2, Y_1, Y_2), the former is integrated with the latter. As a result, *9* out of *21* nodes remain as the characteristic items. It can be analyzed that (X_1, Y_3) is an efficient combination, no combination can become characteristic by adding to (X_2, Y_3), and (X_1, X_2, Y_1, Y_2) is the second most efficient combination.

The bottom figure of Fig. 2 is the result of step 4. Analysts can see that (X_1, Y_3), (X_1, X_2, Y_1, Y_2) and (X_2, Y_1, Y_2) are efficient, in this order, and that the latter two combinations are subsequent efficient items that would not be found by the conventional DEA. Analysts can also see that because the efficiency values of (X_2, Y_1, Y_2) and (X_1, X_2, Y_1) are improved by combing (X_2, Y_1) with the rest of the items, even though the efficiency of (X_2, Y_1) is low, (X_2, Y_1) are key items (as described in section 2.2). As the above description shows, the I/O efficiency structures visualized with the proposed framework allows analysts to determine the items that are characteristics of the DMUs and evaluate their level of efficiency.

4 Numerical Experiment

To examine the utility of the proposed framework, numerical experiments with sample datasets (eighteen DMUs, two input items and four output items) are carried out.

4.1 Considerations Based on the Conventional DEA

A sample dataset is shown in table 1, and the results calculated using the conventional DEA are shown in table 2. As table 2 shows, $DMU_{B, C, I, K, Q}$ are efficient, and the other thirteen DMUs are inefficient. Taking the example of inefficient DMU_R, although weighs are assigned to X_1, X_2, Y_2, and Y_3 and they can be evaluated as strength,

Table. 1. Sample Dataset.

DMU	X_1	X_2	Y_1	Y_2	Y_3	Y_4
A	36	52	2	4	7	5
B	12	31	6	10	11	20
C	25	18	7	10	6	11
D	20	22	8	8	8	8
E	14	31	4	4	9	9
F	23	29	8	12	6	13
G	18	26	4	4	11	15
H	19	31	4	7	8	9
I	11	22	6	8	7	11
J	13	21	3	6	2	6
K	19	21	5	10	6	9
L	19	34	11	4	7	19
M	17	30	7	8	7	10
N	29	36	3	12	7	15
O	22	31	2	7	9	13
P	22	28	6	1	7	15
Q	11	23	9	8	12	19
R	15	24	3	9	9	12

Table. 2. Results by the Conventional DEA.

DMU	Efficiency values	Weights					
		v_1	v_2	u_1	u_2	u_3	u_4
A	0.258	0	0.019	0	0	0.037	0
B	1	0.083	0	0.028	0.083	0	0
C	1	0	0.056	0.141	0.002	0	0
D	0.930	0	0.045	0.115	0.001	0	0
E	0.589	0.071	0	0	0	0.065	0
F	0.941	0.016	0.022	0.014	0.069	0	0
G	0.811	0	0.038	0	0	0.074	0
H	0.581	0.022	0.019	0	0.079	0.004	0
I	1	0.032	0.030	0	0.125	0	0
J	0.720	0.033	0.027	0	0.120	0	0
K	1	0.021	0.029	0.019	0.091	0	0
L	0.827	0	0.029	0.075	0	0	0
M	0.707	0.022	0.021	0.007	0.082	0	0
N	0.753	0.011	0.019	0	0.055	0	0.006
O	0.591	0	0.032	0	0.035	0.039	0
P	0.648	0	0.036	0	0	0	0.043
Q	1	0.091	0	0.111	0	0	0
R	0.951	0.028	0.024	0	0.101	0.005	0

analysts cannot evaluate the efficiency of combinations including Y_1 or Y_4, whose weights are *0*. As is the case of DMU_R, efficient DMU_I has strengths in the combination of X_1, X_2, and Y_2. However, what combinations are characteristics for the efficiency of DMU and how efficient they are cannot be analyzed from results of the conventional DEA.

4.2 Considerations Based on the Framework

We apply the proposed framework to two typical DMUs, visualize the I/O efficient structure and discuss the resulting subsequent efficient items and key items.

Figure 3 is the I/O efficiency structure of DMU_R. Because the dataset has two input and four output items, the number of nodes is $(2^2-1)\times(2^4-1) = 45$. However, the number of nodes visualized decreases to *11* by integrating them, and the efficiency values of (X_1, X_2, Y_2, Y_3), (X_1, X_2, Y_2, Y_4), (X_1, X_2, Y_2), (X_2, Y_2, Y_3), (X_2, Y_2, Y_4), (X_1, Y_2), and (X_2, Y_3) are high in this order. According to the above description, visualizing the I/O efficiency structure helps identify efficient items including the most efficient one that can also be found using the conventional DEA.

However, as the number of out degrees, the number of links from a node to others, from (X_2, Y_2) is *3* and they are linked to nodes whose efficiency values are high – such as (X_1, X_2, Y_2, Y_3), (X_1, X_2, Y_2, Y_4), (X_1, X_2, Y_2), and (X_2, Y_2, Y_3) – it can be found that (X_2,Y_2) plays an important role

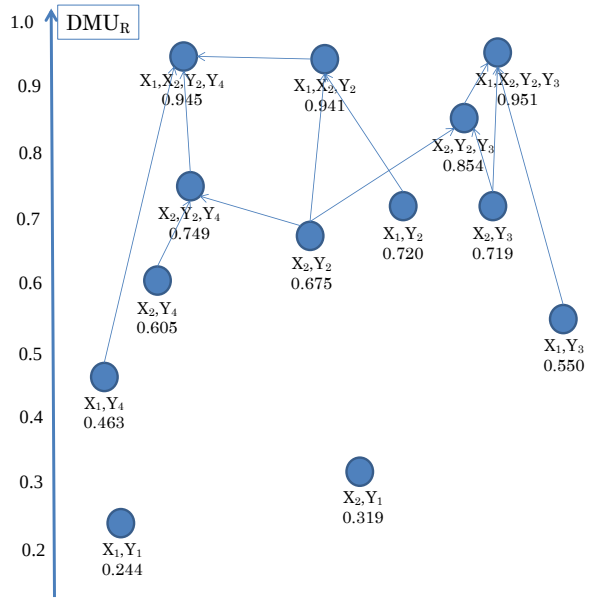


Figure. 3. I/O Efficiency Structure of DMU_R .

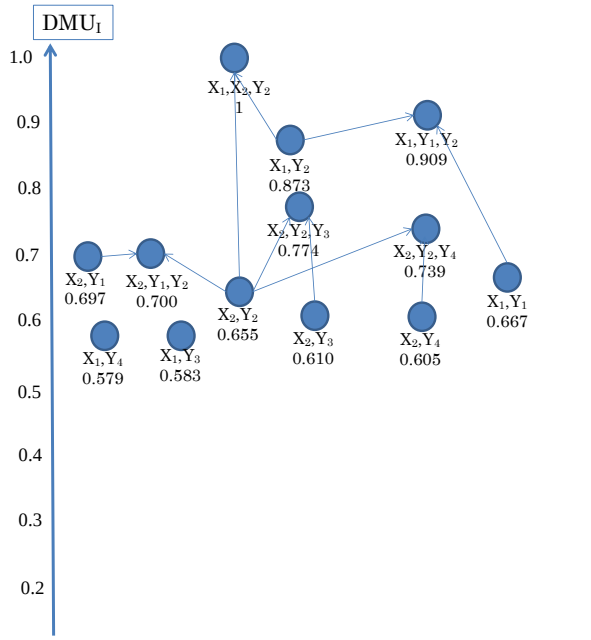


Figure. 4. I/O Efficiency Structure of DMU_J .

for DMU_R and are the key items that improve efficiency values by combining with other items. On the other hand, (X_1, Y_1), and (X_2, Y_1) have no links and do not improve efficiency values by combining. If the combined items are improved, nodes representing them must be visualized.

To summarize them, the I/O efficiency structure shows that the efficiency vales of (X_1, X_2, Y_2, Y_3), (X_1, X_2, Y_2, Y_4), (X_1, X_2, Y_2), and (X_2, Y_2, Y_3) are high in this order, (X_2, Y_2) improve efficiency values in spite of its low efficiency.

Furthermore, (X_1, Y_1) and (X_2, Y_1) cannot improve the efficiency values.

Figure 4 shows the I/O efficiency structure of DMU_I. According to this figure, *13* of *45* nodes are characteristics of DMU_I. According to the result of the conventional DEA, (X_1, X_2, Y_2) are the most efficient items, but the proposed framework allows analysts to evaluate subsequent efficient items as (X_1, X_2, Y_2), (X_1, Y_1, Y_2), (X_1, Y_2), (X_2, Y_2, Y_3), (X_2, Y_2, Y_4), and (X_2, Y_1, Y_2). As is the case of DMU_R, (X_2, Y_2) improves the efficiency of four combinations in spite of its low efficiency value and it can be determined as a key item. On the other hand, the efficiency values of (X_1, Y_4) and (X_1, Y_3) are low and do not improve efficiency values when combined with other items.

5 Conclusion

To evaluate the I/O efficiency structure based on the conventional hierarchical approach, we have proposed a framework composed of four steps. The framework helps analysts to evaluate the characteristic combinations of I/O items for efficiency and the level of their and to find sequent strengths and key items by the visualized I/O efficiency structure. Moreover, numerical experiments with two input and four output items and eighteen DMUs, reveal the utility of *the proposed framework*.

In the future, we will develop a method to integrate more nodes than the proposed framework. As the number of visualized nodes increases with the number of I/O items, with at least *m***s* nodes being visualized, it becomes increasingly difficult to evaluate the I/O efficiency structure. Therefore, it will be necessary to develop a method to decrease the number of nodes. We consider that analysts' intention as important for this purpose.

References

[1] A. Charnes, W. W. Cooper and E. Rhodes: Measuring the efficiency of decision making units, European Journal of Operational Research, Vol.2, pp.429 – 444, 1978

[2] W. W. Cooper, L. M. Seiford and K. Tone: Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA - Solver Software, 2006

[3] K. Inoue, R. Gejima and S. Aoki: Data envelopment analysis for evaluating Japanese Universities, Proceedings of Artificial Life and Robotics, GS12 - 2 pp.110 – 113, 2010

[4] K. Inoue, T. Ichinotsubo and S. Aoki: DEA based Hierarchical Structure and Visualization Method, IEEE International Conference on Fuzzy Systems, pp.1701 – 1704, 2011

[5] J. Doyle and R. Green: Efficiency and cross-efficiency in DEA: Derivations, Meanings and Uses, Journal of Operational Research Society, Vol.45, pp.567-578, 1994

[6] S. Sun and W. M. Lu: A Cross-Efficiency Profiling for Increasing Discrimination in Data Envelopment Analysis. INFOR. Information systems and operational research, Vol.43, pp.51-60, 2005

[7] W. W. Cooper, L. Shanling, L. M. Seiford, K. Tone, R. M. Thall and J. Zhu: Sensitivity and Stability in DEA: Some Recent Developments, Journal of Productivity Analysis, Vol.15, pp.217-246, 2001

[8] A. Charnes, S. Haag, P. Jaska and J.Semple: Sensitivity of Efficiency Calculation in the Additive Model of Data Envelopment Analysis, International Journal of System Science 23, pp.789-798, 1992

[9] T.-Y. Chen and L.-H. Chen: DEA performance evaluation based on BSC indicators incorporated: The case of semiconductor industry, International Journal of Productivity and Performance Management, Vol. 56 Iss: 4, pp.335 – 357, 2007

[10] Y. Yamada, T. Matsui and M. Sugiyama: An Inefficiency Measurement Method for Management System, Journal of the Operations Research Society of Japan, Vol.37, No.2 pp.158-168, 1994

[11] N. Adler, L. Friedman and Z. Sinuany-Sun: Review of ranking methods in the data envelopment analysis context, European Journal of Operational Research 140, 249-265, 2002

[12] D.-Y. Huang and C.-M. Tsou: On some methods for performance ranking and correspondence analysis in the DEA context, European Journal of Operational Research 203, 771-783, 2010

Appendix Data Envelopment Analysis (DEA)

DEA, which was developed by A.Charnes and W.W.Cooper in 1978, is one of the evaluation methods and has been applied to various fields such as data mining in computer science [1]. In DEA, DMUs are assumed as production functions, which produce multiple outputs from multiple inputs and evaluation values are calculated as the ratio of them by comparing DMUs relatively.

Assuming that the number of DMUs is n, the number of input items is m and the number of output items is s, the input values and output values of k th DMU(DMU_k) are expressed as $x_{1k},\ldots, x_{mk}$ and $y_{1k},\ldots, y_{sk}$. The efficiency value of DMU_k is calculated by following Linear Programing (LP) (1). v_i and u_r are weights assigned to the i th input item and the r th output item:

$$
\begin{aligned}
&\max \sum_{r=1}^{s} u_r y_{rk} \\
&s.t. \sum_{i=1}^{m} v_i x_{ik} = 1 \\
&-\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} \le 0 \quad (j=1,\ldots,n) \\
&v_i \ge 0 \quad (i=1,\ldots,m) \\
&u_r \ge 0 \quad (r=1,\ldots,s)
\end{aligned}
\qquad \ldots \text{(A1)}
$$

In DEA, the sum of products of input values and their weights (virtual inputs) is fixed to one, and the objective function is to maximize the sum of products of the output values and their weights (virtual outputs). The second constraint implies that the efficiency values of other DMUs are less than one. In other words, each DMU adopts weights that maximize the efficiency value under the condition that the efficiency values of other DMUs do not exceed one. Therefore, the strengths of DMUs can be evaluated by the weights calculated bia DEA.

$$
\begin{aligned}
&\min \ \theta \\
&\text{s.t.} -\sum_{j=1}^{n} \lambda_j x_{ij} + \theta x_{ik} \ge 0 \quad (\text{i}=1,\ldots,\text{m}) \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - y_{rk} \ge 0 \quad (\text{r}=1,\ldots,\text{s}) \\
&\lambda_j \ge 0 \quad (\text{j}=1,\ldots,\text{n})
\end{aligned}
\qquad \ldots \text{(A2)}
$$

The dual problem of (A1) is described as following LP (A2) with θ and λ_j as dual variables. As (A2) is the dual problem of (A1), (A2) has the same optimal solutions as (A1), and θ is the efficiency value. The constraints of (A2) imply that DMU_k is included in a convex combination of all DMUs and the objective function is to minimize the sum of input values under the condition of maintaining the present output values. That is, if DMU_k is inefficient, its efficiency value is calculated by comparing the ratio to the point on the frontier composed of DMUs whose λ_j are positive.

As described above, DEA has two features. One concern the weights assigned to the input and output items. In DEA, efficiency values are calculated by the ratio of virtual input and output values which are sums of products of input and output values and their weights. The weights are different for each DMU and analysts can evaluate strengths from them as DMU can adopt weights that maximize the efficiency value. The other feature concerns the efficiency values. In DEA, DMUs of best practice are equally efficient and their efficiency values are equal to one. The efficiency values of other inefficient DMUs are calculated based on comparison of present DMU

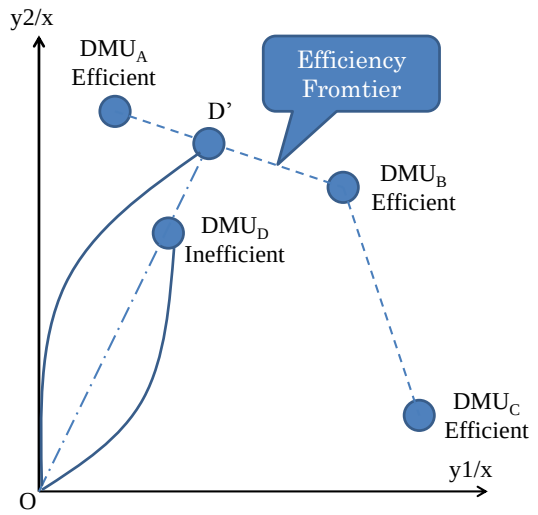


Figure. The image of DEA

and a point on the efficiency frontier.

The above figure is the image of DEA with one input item and two output items. $DMU_{A,B,C}$ are efficient and DMU_D is inefficient. The efficiency values of $DMU_{A,B,C}$ are one equally and that of DMU_D is calculated as OD/OD'.

Hybrid DEA Model Based on Hierarchical Approach

Kenji Fukuoka† **Yukie Majima†** **Kazushige Inoue†** **Shingo Aoki‡**

†Osaka Prefecture University, Japan
{fukuoka, inoue}@mis.cs.osakafu-u.ac.jp
majima@cs.osakafu-u.ac.jp
‡Hiroshima Institute of Technology, Japan
s.aoki.sm@it-hiroshima.ac.jp

Abstract

This paper proposes HH-DEA (Hybrid Hierarchical DEA) model composed by DEA (Data Envelopment Analysis) and IDEA (Inverse DEA) based on hierarchical approach. HH-DEA model can analyze the improvement direction of DMU (target DMU) from two viewpoints, strength and weakness, by forming multi efficient frontiers and inefficient frontiers. As the result, proposed method is able to find the change of inefficiency by IDEA compared with traditional method. The advantages of proposed method are shown by numerical experiments with sample dataset, fifteen DMUs, two input and output items.

Keywords: Data Envelopment Analysis, Operation Research, Hierarchical Approach, Inverse DEA

1 Introduction

Data Envelopment Analysis (DEA) is one of methods to analyze management efficiency of Decision Making Unit (DMU) [1], such as private company, hospital or so [2]. It is applied in various fields such as data mining. In DEA, each DMU is characterized by multi-input and multi output items. The following two analyses are mainly performed, (1) efficiency value which shows efficiency in consideration of strengths of DMU is calculated by the rate of an output item to an input item, (2) target DMU, improvement direction to only inefficient DMU. The target DMU is calculated based on the efficient frontier formed of efficient DMU. Inverse DEA (IDEA) was proposed by Yamada [3] to analyze DMU from inefficiency in consideration of their weakenesses rather than their strengths. This method calculated inefficiency value in consideration of weaknesses of DMU. Thus, in DEA, not only efficiency in consideration of strengths but also inefficiency in consideration of weaknesses is evaluated. Moreover, an hybrid method of them was proposed [4].

There is hierarchical approach as research on the amount of improvements for becoming efficient [5]. This approach allows to analyze a stepwise improvement direction using not only efficient DMU but using inefficient DMU. This method is used for searching stepwise target DMUs and improvement direction by dividing an efficient frontier. Here, it allowed to discover which DMU each DMU should aim because they became efficient.

While this approach guarantees that target DMU is in the class of a higher rank in respect of efficiency rather than self, in respect of inefficiency like IDEA to the traditional DEA, it is not possible to take into consideration what kind of situation DMU was in. Then, whether an improvement target was the extreme proposal that both efficiency and inefficiency go up, whether an improvement target was the robust proposal that the former go up and the latter go down, and the efficiency and inefficiency over an improvement target were not possible to be analyzed. Moreover, information, including these efficiency, inefficiency, etc., are important for a decision maker because there are many improvement targets along with it when there are many classes of a frontier.

Therefore, this paper proposes HH-DEA (Hybrid Hierarchical DEA) model composed by DEA and IDEA based on hierarchical approach. HH-DEA model can analyze the stepwise target DMU from two viewpoints, strong and weak points. There are three analysis steps. Step1: Multi efficient frontiers are formed and efficient value and reference set of DMU are calculated based on each efficient frontier as with traditional

hierarchical approach. Step2: Multi inefficient frontiers are formed and inefficient value of DMU is calculated based on each inefficient frontier. Step3: The figure is illustrated by these efficiency value, reference set and inefficient value. An analyst is able to set up target DMU in consideration of these strength and weakness of DMU.

A numerical experiment is performed with sample dataset, fifteen DMUs, two input and output items to show the advantages of proposed method.

In this paper, Chapter 2 will explain DEA. Chapter 3 will describe framework by hierarchical approach and the problem. Chapter 4 will explain the proposal method and Chapter 5 will show the result in numerical experiments and will describe the consideration. Chapter 6 will conclude this paper.

2 Data Envelopment Analysis

Section 1 describes the outline of DEA and the way to calculate efficiency value and reference set. Section 2 describes the outline of IDEA and the way to inefficiency value.

2.1 DEA

DEA was proposed by A. Charnes et al. in 1978 as a method for management analysis. It is applied in various fields such as data mining.

DMU is characterized by input and output items and its efficiency is evaluated in consideration of strengths. This efficiency is evaluated by efficiency value. Then, efficient DMU has efficiency value as 1, inefficient DMU has value lower than 1. Efficiency value is calculated based on line, called efficiency frontier. Efficiency value of inefficient DMU is calculated based on line called efficiency frontier, which is formed by efficient DMUs. Then, an inefficient DMU is estimated to be inefficient because of some efficient DMUs. They are called reference set.

Assuming that there are *n DMUs* (DMU_1, DMU_2, …, DMU_k, …, DMU_n), and DMU_k is characterized by *m* inputs and *s* outputs. Then DMU_k has input expressed as $x_{1k},\ldots,x_{mk}$ and output expressed as $y_{1k},\ldots, y_{sk}$. Here the efficiency value and reference is shown by calculating following linear algorithm.

$$
\begin{aligned}
\max \quad & \theta_k \\
s.t. \quad & \sum_{j=1}^{n} x_{ij}\lambda_{jk} \le x_{ik} \quad (i=1,2,\ldots,m) \\
& \sum_{j=1}^{n} y_{rj}\lambda_{jk} \ge \theta_k y_{rk} \quad (r=1,2,\ldots,s) \\
& \theta : free
\end{aligned}
\tag{1}
$$

The optimal solution $(\theta^*, \lambda_1^*.., \lambda_n^*)$ is calculated by the formula (1). When $\theta_k^* \le 1$, an efficiency value is expressed as θ_k^* at this time. When $\theta_k^* = 1$. DMU_k is evaluated to be efficient. And when $\theta_k^* \le 1$, DMU_k is evaluated to be inefficient. A reference set is expressed as the set shown by $(\theta^*, \lambda_1^*.., \lambda_n^*)$.

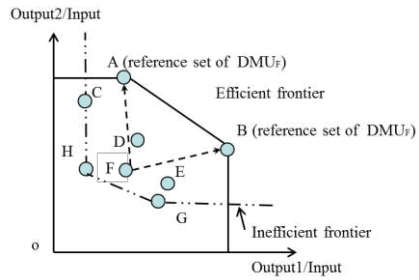


Fig.1 Example of DEA and IDEA

2.2 Inverse DEA

In order to measure the efficiency of DMU, the method of focusing on the strong and weak points is also considered. IDEA is method for evaluating inefficiency in consideration of weakness, as opposed to DEA, which evaluates efficiency in consideration of the strength of DMU. This inefficiency is evaluated by inefficiency value. Then IDEA-inefficient DMU has inefficiency value as 1, IDEA-efficient DMU has value lower than 1. Moreover, IDEA is added like IDEA-efficient and IDEA-inefficient DMU in order to distinguish from efficient and inefficient DMU in DEA as expression.

Next, the way to calculate inefficiency value.is described. In IDEA, the assumption about DMU or an input-and-output item is the same as that of DEA.

The optimal solution θ_k^* calculated by the formula (2). There is $\theta_k^* \le 1$ and an inefficiency value is expressed as θ_k^* at this time. When $\theta_k^* = 1$. DMU_k is evaluated to be IDEA-inefficient. And when $\theta_k^* \le 1$, DMU_k is evaluated to be

IDEA-efficient.

$$\begin{aligned} \max \quad & \theta_k \\ s.t. \quad & \sum_{j=1}^{n} x_{ij}\lambda_{jk} \le x_{ik} \quad (i=1,2,...,m) \\ & \sum_{j=1}^{n} y_{rj}\lambda_{jk} \ge \theta_k y_{rk} \quad (r=1,2,...,s) \\ & \theta : free \end{aligned} \tag{2}$$

3 Hierarchical approach and its problem

3.1 Hierarchical approach

When DMU is evaluated inefficient, not only the efficiency value of its DMU is calculated but also the target DMUs, improvement direction, to become efficient can be calculated as the feature of DEA. However, these target DMUs may be unreal because they are obtained by comparison with a first place group of the business world. The stepwise target DMUs to improve the efficiency was not calculated for inefficient DMU in traditional DEA model. Therefore, in the previous work, the method allowing to analyze a stepwise improvement direction by using inefficient DMU which was not taken into consideration when setting up target DMUs. The improvement direction is proposed by dividing an efficient frontier by class. The sensitivity analysis [6] in DEA is used to divide by an efficient frontier. This method forms multi frontiers by gradually removing efficient DMU whose efficiency value is 1. This method has two analysis steps: division of efficient frontier by class, and visualization of the improvement direction.

How to divide an efficient frontier by class will be explained. Firstly, the efficiency value and reference set of each DMU are calculated by traditional DEA. Then, DMUs that have efficiency value 1 form frontier_1 and belong to the first place group. Next, the efficiency value and reference set are calculated again using dataset except DMU of first place group. Here, DMUs which have efficiency value 1 form new frontier_2 and belong to the second place group. Similarly, multi frontiers are formed by removing DMUs which are evaluated as efficiency value 1 by traditional DEA. Fig.2 shows an example of hierarchical approach. Firstly, the efficiency value and reference set of each DMU, $DMU_{A,..,H}$, are calculated by traditional DEA. Then, $DMU_{A,B}$ which have efficiency value 1 form frontier_1 and belong to first place group. Next, the data set $DMU_{C,..,H}$ except $DMU_{A,B}$ is created. The efficiency value and reference set are calculated again using new dataset by traditional DEA. Then, DMUs which have efficiency value 1 form frontier_2 and belong to second place group. This procedure is performed until it becomes impossible to create new data set. In the case of Fig. 2, this procedure is performed until the frontier_3 is created.

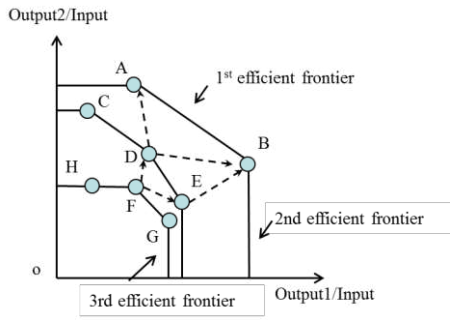


Fig.2 Example of DEA hierarchical approach

How to visualize the improvement direction will be explained. Some reference sets are used for visualization. When an DMU which belongs to frontier_k refers to some DMU(s) which belong to frontier_(k-1), some arrows are pulled from the DMU to them. Thus, the stepwise improvement direction can be analyzed from the figure visualized by the reference set calculated on each frontier.

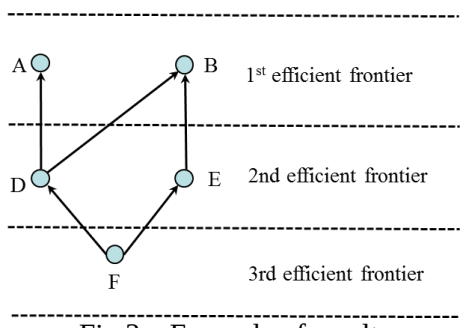


Fig.3 Example of result by DEA hierarchical approach

Fig. 3 shows the result of visualization focusing on stepwise target DMUs and improvement direction of DMU_F. Then, DMU_F which belongs to frontier_3 refers to $DMU_{D,E}$ which belong to frontier_2. It is analyzed that these DMU are target of DMU_F. Similarly, DMU_D refers to $DMU_{A,B}$ and DMU_E refers to DMU_B from the relation between frontier_1 and frontier_2. It

is analyzed that these DMU_A are target of $DMU_{D,E}$ and DMU_B are target of $DMU_{E.}$ Although the stepwise target DMUs and improvement direction was not calculated, it is possible to show these by hierarchical approach.

3.2 Problem

In traditional DEA model, these target DMUs may be unreal because they are obtained by comparison with a first place group of the business world. The traditional hierarchical approach solved the problem by hierarchically forming an efficient frontier and calculating the realizable stepwise improvement direction. This approach guarantees that target DMU is in the class of a higher rank in respect of efficiency rather than itself. However, in respect of inefficiency like IDEA to the traditional DEA, it was not possible to take into consideration what kind of situation DMU was in. Then, whether an improvement target was the extreme proposal that both efficiency and inefficiency go up, whether an improvement target was the robust proposal that the former go up and the latter go down, the efficiency and inefficiency over an improvement target were not possible to be analyzed. In Fig. 2, which is an example of DEA hierarchical approach, $DMU_{D,E}$ is target of $DMU_{F.}$ Then, these DMU is not judged whether it is an above-mentioned extreme or robust proposal because their efficiency and inefficiency are not shown. Moreover, this information, including these efficiency, inefficiency, etc., are important for a decision maker because there are many improvement targets along with it when there are many classes of a frontier.

4 HH-DEA (Hybrid Hierarchical DEA)

This paper proposes HH-DEA (Hybrid Hierarchical DEA) model composed by DEA and IDEA based on hierarchical approach. HH-DEA model can analyze the stepwise target DMU from two viewpoints, strong and weak points by forming multi frontiers. This model has three analysis steps, these are as follows: (Step1) DEA based on hierarchical approach introduced in Chapter 3, (Step2) IDEA based on hierarchical approach and (Step3) Visualization by efficiency value and reference set calculated in Step1 and inefficiency value calculated in Step2.

In Step1, in order to analyze stepwise target DMUs and improvement direction in consideration of a strong points, the traditional hierarchical approach in Chapter 3 is used. This approach is able to set up target DMUs and the improvement direction by dividing an efficient frontier by class. Multi frontiers are formed by gradually removing efficient DMU whose efficiency value is 1 to divide frontier, and the efficiency value and a reference set of DMU are calculated on the basis of each frontier.

In Step2, in order to add the evaluation in consideration of a weak point, IDEA analysis based on traditional hierarchical approach is performed. How to divide an inefficient frontier by class is explained and is similar to the traditional method. Firstly, the inefficiency value of each DMU is calculated by traditional IDEA. Then, DMUs which have inefficiency value 1 form non-frontier_1 and belong to the first place group. Next, the inefficiency value is calculated again using dataset except DMU of the first place group. Here, DMUs which have inefficiency value 1 newly form non-frontier_2 and belong to the second place group. Similarly, multi non-frontiers are formed by removing DMUs which are evaluated as inefficiency value 1 by traditional IDEA. Fig.4 shows the example of this analysis. Firstly, the inefficiency value of each DMU, $DMU_{A,..,H}$, is calculated by traditional IDEA. Then, $DMU_{C,H,G}$ which have inefficiency value 1 form non-frontier_1 and belong to the first place group. Next, the data set except $DMU_{C,H,G}$ is created. The inefficiency value is calculated again using new dataset by traditional IDEA. Then, DMUs which have inefficiency value 1 form non-frontier_2 and belong to the second place group. This procedure is performed until it becomes impossible to create a new data set. In the case of Fig. 4, this procedure is performed until the non-frontier_3 creates.

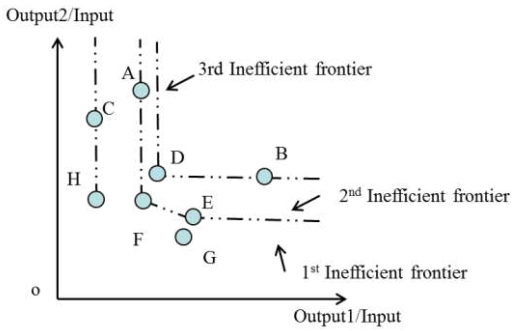


Fig.4 Example of result by Step2, IDEA hierarchical approach

Finally, in Step3, stepwise target DMUs and improvement direction are visualized by efficiency value and reference set calculated in Step1 and inefficiency value calculated in Step2. How to visualize will be explained. Firstly, the vertical axis of a figure is set as the class number of non-frontier when inefficiency value of the DMU is 1 and the horizontal axis of a figure is set as the class number of frontier when efficiency value of the DMU is 1. Next, all DMU is plotted in the figure. When an DMU which belongs to frontier_k refers to some DMU(s) which belong to frontier_(k-1), some arrows are pulled from the DMU to them. Thus, we can analyze the stepwise target DMU from two viewpoints, strong and weak points. The stepwise target DMU from two viewpoints, strong and weak points, can be analyzed by creating figure. Fig.5 show the result of visualization.

Table1 Example of result by Step1 and 2 of proposed model

DMU	Number of	
	Effcient Frontier	InEfficient Frontier
A	1	2
B	1	3
C	2	1
D	2	3
E	2	2
F	3	2
G	3	1
H	3	1

In Fig. 5, the efficiency of DMU which exists in the left in a horizontal axis in consideration of a strong points is high. The inefficiency of DMU which exists downward in a vertical axis in consideration of a weak point is also high. Thus, the DMU in the upper left can be analyzed to be robust and have many strong points and few weak points. The DMU in the lower left can be analyzed to be peculiar and have many strong points and weak points. When focusing on DMU_F, DMU_D is analyzed to be target whose efficiency is higher and inefficiency is lower than DMU_F. And DMU_E is analyzed to be target whose efficiency is higher than DMU_F and inefficiency is same. In traditional hierarchical approach, it was unclear how the inefficiency in consideration of weakness would change. However, HH-DEA model can analyze the stepwise target DMU from two viewpoints, strong and weak point, by adding evaluation of IDEA.

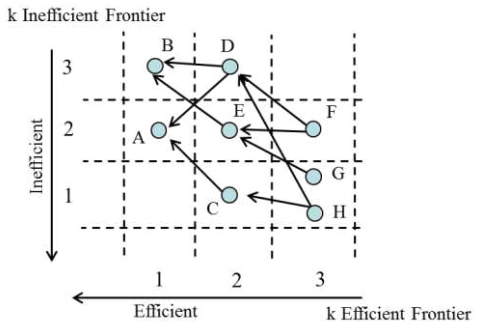


Fig.5 Example of result by Step3, visualization

5 Numerical experiment

A numerical experiment is performed with sample dataset, fifteen DMUs, two input and output items to show the advantages of proposed method. The traditional hierarchical approach and the proposal method are applied to sample data, and each result is compared.

Figs. 6 and 7 show the result of the traditional hierarchical approach and the proposal method respectively, and the arrow in the figures shows the reference place, target DMU. An efficiency value, a reference set and an inefficiency value are calculated in order to visualize target DMUs, and improvement direction are indicated in appendix. Some DMUs which are not target of DMU1 is not plotted to make a figure legible .

Table2 Sample data

DMU	Input		Output	
	X1	X2	Y1	Y2
DMU1	8	14	4	16
DMU2	19	3	17	7
DMU3	10	15	12	20
DMU4	5	7	12	2
DMU5	4	16	3	20
DMU6	16	4	15	8
DMU7	12	16	7	7
DMU8	3	20	4	16
DMU9	20	2	4	1
DMU10	4	2	10	20
DMU11	11	1	8	9
DMU12	8	19	14	7
DMU13	15	18	13	3
DMU14	2	10	2	12
DMU15	13	7	11	12

5.1 Consideration by the traditional hierarchical approach

Fig. 6 shows the result of the traditional hierarchical approach. Here, Fig.6 shows that DMU1 aims at other inefficient DMU15 and 5, and then that DMU1 aims at DMU10 and 14. Thus, the stepwise target DMU for improving the efficiency in consideration of strength of DMU is analyzed. However, these DMU is not judged

whether it is an above-mentioned extreme or robust proposal because their efficiency and inefficiency are not shown.

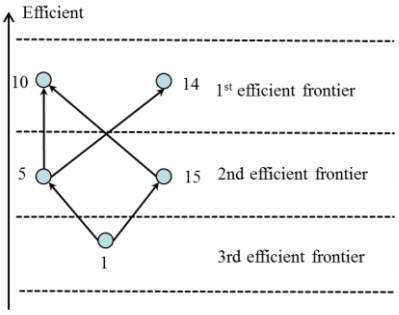


Fig.6 Visualization by the traditional hierarchical approach

5.2 Consideration by proposed method

The stepwise improvement direction of DMU is analyzed from two viewpoints, strong and weak points, which were not able to be analyzed with the visualized figure.

Fig. 7 shows the result of proposal method. Since DMU10 is located in the upper left of the figure, it is robust while DMU11 has a weak point compared with DMU10, and it has analyzed that it is peculiar.

Table3 Result by Step1 and 2 of proposed model

DMU	Number of	
	Efficient Frontier	Inefficient Frontier
DMU1	3	2
DMU2	2	2
DMU3	2	3
DMU4	2	2
DMU5	2	1
DMU6	2	3
DMU7	4	1
DMU8	2	1
DMU9	3	1
DMU10	1	4
DMU11	1	2
DMU12	3	2
DMU13	4	1
DMU14	1	2
DMU15	2	3

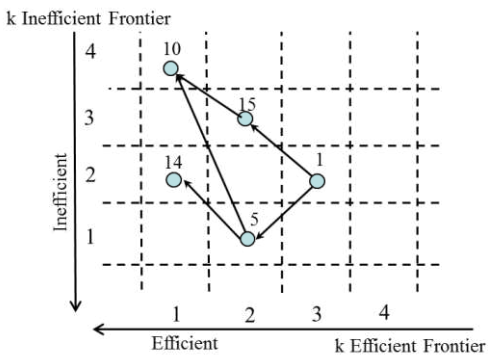


Fig.7 Visualization by proposed method (Step3)

6 Summary

This paper proposes HH-DEA (Hybrid Hierarchical DEA) model composed by DEA and IDEA based on hierarchical approach. There are the following three analysis steps. Step1: Multi efficient frontiers are formed and efficient value and reference set of DMU are calculated based on each efficient frontier as with traditional hierarchical approach, Step2: Multi inefficient frontiers are formed and inefficient value of DMU is calculated based on each inefficient frontier, Step3: the figure is illustrated by these efficient value, reference set and inefficient value. An analyst is able to set up target DMU based on these DMU's strength and weakness. Result of numerical experiment shows the advantages of the proposed method.

As future works, a new method considering the width between each frontier will be proposed.

References

[1] A. Charnes, W. W. Cooper, and E. Rhodes : "Measuring Efficiency of Decision Making Units," European Journal of Operational Research, Vol.2, No.6, pp.429-444 (1978).

[2] W. W. Cooper, L. M. Seiford, and K. Tone : "Data Envelopment Analysis A comprehensive Text with Models, Applications, References and DEA"Solver Software Second Edition," Springer (2007).

[3] Yamada Y, Matsui T, Sugiyama M. "An inefficiency measurement method for management systems". Journal of Operations Research Society of Japan, Vol.3, pp158–168 (1994).

[4] Ukita, M., Yanagawa, Y., and Miyazaki, S: "An Improvement for a Business Efficiency Analysis Method Using DEA and Inverted DEA", Japan Industrial Management Association, Vol.60, No.3, pp.119-127 (2009).

[5] Sato, Y., Inoue, K., and Aoki, S: "Numerical study of search method for DEA-improvement direction based on hierarchical approach", The Institute of Electrical Engineers of Japan, pp.1038-1043 (2007).

[6] N. Hibiki, and T. Sueyoshi: "DEA sensitivity analysis by changing a reference set: regional contribution to Japanese industrial develoment", The International Journal of Management Science, Vol.27, pp139-153 (1999).

Markov Random Fields based Estimation of Distribution Algorithm and Application of Sharable Resource Constrained Scheduling Problem

X. Hao†* J. Tian† H. W. Lin‡ T. Murata†

† Graduate School of Information Production and Systems,
Waseda University, Kita-Kyushu, Japan
‡ Harbin Institute of Technology,
Shenzhen Graduate School Shenzhen, China
haoxc@ruri.waseda.jp

Abstract

During the past several years, a large number of studies have been conducted in the area of sharable resource constrained scheduling problems. Intelligent manufacturing planning and scheduling based on meta-heuristic methods, such as the simulated annealing and particle swarm optimization, have become common tools for finding satisfactory solutions within reasonable computational times in real settings. However, only a few studies have analyzed the effects of interdependent relations during group decision-making activities. This paper proposes a Markov network based estimation of distribution algorithm (MNEDA) to overcome these challenges. We present an empirical validation of MNEDA on the semiconductor final test scheduling problem as an example of resource constraint scheduling problems.

Keywords: Markov Random Fields, Estimation of Distribution Algorithm, Sharable Resource constrained Scheduling Problems.

1 Introduction

Resource constrained scheduling problems (RCSP) have attracted an ever growing attention in recent years both from science and practice. It is concerned with single-item or small batch production where scarce resources have to be met when scheduling dependent activities over time[1]. Particularly, flexible manufacturing systems (FMS) allow the dynamic configuration of resources to process distinct products. It is possible to assign each of these products to more than one configuration of unrelated resource with various efficiencies. Multipurpose resources can perform a wide range of tasks, which allows schedulers to concentrate the workloads among these resources to improve the usage rate and reduce resource requirements. On the other hand, due to the high capital expenditure in capacity investment, there is limited availability of this common resource. Therefore at any single time point the amount of this resource consumed cannot exceed its maximum availability. These types of resource refer to a renewable resources[1], [2], [3]. A possible negative effect is the reduction of production effectiveness because of the time wasted when a machine performs changeover and configuration to accommodate for the next job. Thus, it is vital to arrange a production schedule that simultaneously considers the values of multiple resources to respond to production requirements rapidly and effectively[3], [4].

Recently, shared resource constrained scheduling problem (SRCSP) as a scenario of RCSP has been widely researched. Intelligent manufacturing planning and scheduling based on meta-heuristics, such as genetic algorithms (GAs), simulated annealing (SA), and ant colony optimization (ACO) have become common tools for finding satisfactory solutions within reasonable computational times in a real setting. Ballestin and Trautmann present a population-based iterated local search method for a resource constrained project scheduling problem, where tasks have prescribed completion times and deviation from these times across all tasks is required to be minimized[5]. Wu and Chien propose a mixed-integer linear programming (MIP) model for semiconductor final testing of job scheduling problem(SFTSP) and dynamic testing of machine configurations and develop an assignment algorithm[6]. Singh and Ernst discuss a problem in

which jobs do not have hard deadlines and need to be scheduled on multiple machines. The objective is to minimize the total weighted tardiness, subject to release dates, precedence and resource constraints. For this problem, they compare a Lagrangian relaxation-based approach with simulated annealing[7]. Thiruvady et al consider shared SRCSP in the mining industry and present hybrid of ACO incorporating bean search and constraint programming.

Although previous studies have shown that the traditional meta-heuristic is a powerful optimization tool that performs effectively in real applications, three primary reasons suggest that the algorithm is not entirely adequate for solving SRCSP. For SRCSP, strong interdependent relationship exists between the resource (process plan) of jobs and the operation sequence of jobs. However, interaction among decision variables of various hierarchy levels is inadequately analyzed in previous studies.

This paper proposes a Markov network based estimation of distribution algorithm (MNEDA) to solve SRCSP. MNEDA describes an EDA based on local Markov property, which estimates conditional probabilities in terms of neighboring nodes in the undirected graph and samples from them using a Gibbs sampler.

The remainder of this paper is organized as follows: Section 2 presents the proposed MNEDA approach in detail; Section 3 provides a review of the SRCSP and presents experimental comparisons that apply the MNEDA approach for analyzing and solving several SFTSF; and finally, Section 4 offers a conclusion.

2 Proposed Markov Random Fields based EDA

Estimation of Distribution Algorithms (EDAs) is a class of population based optimization algorithm that extracts statistical information from the population of solutions and uses it to generate new solutions[8], [9].

Estimation of the probability distribution lies in the very heart of an EDA. Its effectiveness largely depends on how well it estimates and samples the probability distribution. Probabilistic graphical models, which is used to represent the interaction among the discrete variables, have been attracted an ever growing attention in recent years.[10], [11], [12], [13]. Probabilistic graphical models can be mainly categorized into two groups: Directed models (Bayesian networks) and Undirected models (Markov networks/Markov Random Fields).

2.1 Markov networks

A Markov network is a pair (G, Ψ); where G is the structure and the Ψ are the parameter set of network. G is an undirected graph where each node corresponds to a random variable in the modelled data set and each edge corresponds to conditional dependencies between variables. However, unlike the Bayesian networks, the edges in Markov networks are undirected. Here, the relationship between two nodes should be seen as a neighborhood relationship, rather than a parenthood relationship. We use N to define a neighborhood system on G, where each N_i is the set of nodes neighboring to a node X_i.

A Markov network is characterized in term of a neighborhood relationship between variables by its local Markov property, that is, a variable in Markov network is conditionally independent of all other variables given its neighbors. It states that the conditional probability of a node X_i given the rest of the variables can be completely defined in terms of the conditional probability of the node given its neighboring states N_i. In term of probability, it can be written as

```
procedure: Markov network-Based EDA routine
        input:   problem data, parameters ( M, N)
       output:   the best solution S_best
begin
    initialzation:
            t ← 0;
     step1: initialize population Pop(0) size M randomly;
    Optimization:
    while( not meeting termination criterion)
        step2: Select set D from Pop(t) consisiting
               of N solutions, where N < M.
        step3: Estimate the Structure of Markov network
               from D.
        step4: Estimate local Markov the conditional
               probabilities p(x_i|N_i), for each variable X_i,
               samples candinates based on p(x_i|N_i).
        step5: evaluate candinates and update the
               best solution S_best.
        step6: P(t+1) = adjust(P(t), candinates).
               t ←t+1.
    end;
end;
```

Fig1. Psedo-code for MNEDA

Step3: Estimating the structure of Markov network

1. Generate a matrix of mutual information which estimates the mutual information for each pair of variables in the solution. The mutual information between two random variable, *A* and B is given by
$$MI(A;B)=\sum_{b\in B}\sum_{a\in A}p(a,b)\log\left(\frac{p(a,b)}{p(a)p(b)}\right)$$
2. Create an edge between two variables, if the mutual information between them is higher than the threshold $MT = AVG(MI)*\alpha$, where $AVG(MI)$ is the average of the elements of the MI matrix and coefficient α is the significance parameter that controls the density of the edges.
3. If the number of neighbors to a variable is higher than the maximum number, *MN*, allowed, only keep the *MN* neighbors that have the highest mutual information.

Fig2. Psedo-code for Estimating the structure

$$p(x_i \mid x-\{x_i\}) = p(x_i \mid N_i) \tag{2}$$

As the Markov properties of an arbitrary probability distribution can be difficult to establish, a commonly used class of Markov random fields are those that can be factorized according to the cliques of the graph. The joint probability distribution can be written as

$$p(\mathrm{x}) = \frac{1}{Z}\prod_{i=1}^{m}\psi_i(c_i) \tag{3}$$

where, $\psi_i(c_i)$ is a potential function on clique $c_i \in X$, and m is the number of cliques in the structure G. the normalizing constant $Z=\sum_{x\in\Omega}\prod_{i=1}^{m}\psi_i(c_i)$ Z known as the *partition function* which ensures that $\sum_{x\in\Omega}p(x)=1$.

2.2 Local Markov Property based EDA

Nowadays, most of the Markov network based EDAs factorize the joint probability distribution represented in equation(3) in term of the cliques in the undirected graph and sample it to generate new solutions[9], [14], [15]. Furthermore, they suffer from curse of dimensionality even for small cliques. It leads to complex approximation involving cliques and the performance high depends on the definition potential functions. In this paper, Local Markov property of Markov network in EDA (MNEDA) is employed as an alternative to the joint probability distribution.

MNEDA starts by generating a population solutions. A set of solutions (promising data) is selected from the population using a selection method, and the promising data is used to learn the structure of Markov network. Thereafter, the conditional probabilities defined by the local Markov property are then estimated from the promising data. Finally, the new candidate solutions are incorporated into a solution pool, which keeps these individuals contribute to the makeup of promising data. The iteration will continue until the predefined termination criteria is met. The pseudo-code for MNEDA is presented in Fig. 1.

A number of different approaches can be used to estimate an undirected structure in term of one method of mutual information, L1-regularization etc. In order to avoid high order tests for an excessively complex network, the pairs with the mutual information are used to estimate the structure. Fig. 2 presents the implemented Markov network structure learning algorithm. The MN parameter has a directed effect on the balance between the accuracy and complexity of the model. The higher the MN value is, the more number of neighbors needed to be used to represent the conditional probability.

Once the structure of Markov network has been estimated, the new candidate solutions are

Step4: Gibbs Sampling based on local property of Markov network

begin
  step1: Generate a candidate solution $x = (x_1, x_2, \ldots, x_n)$ at random.
  For i = 1 to n do
    step2: Choose a variable x_i from the solution x.
    step3: Estimate the Structure of Markov network from *D*.
    step4: Using selected set of solution *D*, estimate the conditional probability $p(x_i|N_i)$, for each value x_i of the variable X_i. as Gibbs probability:
$$p(x_i \mid N_i) = \frac{e^{p(x_i,N_i)/T}}{\sum_{x_i'\in X_i} e^{p(x_i',N_i)/T}}$$
    step5: sample the conditional probability $p(x_i|N_i)$ to new x_i.
  end;
end;

Fig3. Psedo-code for Gibbs Sampler

sampled from it. Most of the Markov networks do not satisfy the ancestral ordering of variables needed by probabilistic logic sampling, Gibbs sampler, a class of Markov Chain Monte Carlo method, as the sampling method in our algorithm. The pseudo-code is described in Fig. 3.

3 Experiments and Discussion

3.1 Sharable Resource Constrained Scheduling Problem

SRCSP is one of the most complex and common seen scheduling problems. In SRCSP, multiple resources are simultaneously combined and aligned to process a job. Without loss of generality, a set of jobs $j \in \mathbf{J}$ are ready to be scheduled for a short-term planning horizon. Each job is inseparable so that the process characteristics can be preserved for quality assurance and yield enhancement. Furthermore, each job has to be processed through a set of operations $o \in \mathbf{O}_j$. Normally, after an operation starts processing on a particular resource, it has to continue until completion, which is a non-preemption constraint.

Resource types can vary and include dedicated physical equipment, platforms, or interfaces between the equipment and goods, and supporting labor. There are N types of heterogeneous resources $\mathbf{R}^1 = \{R_1^1,...,R_{\nu^1}^1\}, \mathbf{R}^2 = \{R_1^2,...,R_{\nu^2}^2\},..., \quad \mathbf{R}^N = \{R_1^N,...,R_{\nu^N}^N\}$ for machine configuration, each of which contains homogeneous but unrelated resources, i.e. with different performances. Some of them can be extremely expensive and thus only be prepared in limited amount, say Q_r, $r \in \mathbf{R}^n, n \in \{1,2,...,N\}$, according to long-term capacity planning. Thus, the sum of operations overlapping in time demanding any given resource should not exceed the available total amount. Multiple factors, such as maintenance, repairs, breakdowns, and ongoing R&D experiments can influence the availability of resources This study considers static resource availability whereas further research should take stochastic unavailability and dynamic release of multiple resources into account.

On the other hand, machine configuration is dynamic. Different operations may require different machine (configuration) types for processes, whereas more than one machine type can process one operation, which reveals the virtual unrelated parallel machine environment. Different machine types $m = (r^1,...,r^n,...,r^N) \in \mathbf{M}_o, r^n \in \mathbf{R}^n$, with dissimilar processing time, say P_{om}, can perform an operation o. However, it is not possible to configure resources arbitrarily to form a machine type. Consequently, the universal set of all machine types is $\mathbf{M} = \bigcup_{o \in \mathbf{O}_j, j \in \mathbf{J}} \mathbf{M}_o \subseteq \mathbf{R}^1 \times ... \times \mathbf{R}^N$. To disassemble the original machine type l and assemble and calibrate the new machine type m for the incoming operation, sequence-dependent setup time (SDST), denoted by S_{lm}, and is further. Setup activities comprise resources assembly and disassembly, temperature change, software download, and calibration. The setup is separable from the process and is anticipatory so that a setup can start before a job is ready[16]. The reconfiguration decision is centralized control rather than hierarchical one that is interacted with multiple local and autonomous controllers. A poor plan on machine configuration may waste resource availability on redundant setup or, on the other hand, lead to increasing unmet demand.

To examine the practical viability and efficiency of proposed MNEDA, we designed a numerical study for comparison MNEDA needed to be compared with efficient algorithms in the previous work. The proposed MNEDA was compared with wcGA[6] based on a set of simulation data of semiconductor final testing scheduling problems.. Parameters of simulation data were aligned in order to provide a comparative basis. The large scale problem contains 100 jobs and each job has 3 operations. The source contains 3 type tests, 3 type handlers and 4 type Accessories. The quality of resources is (10, 8, 4), (7, 7, 5, 5) and (10, 5, 3) respectively. The experiments were conducted on a personal computer with an Intel Core I5 CPU at 2.8G and 2GB RAM. The parameters and strategies of related algorithms are categorized in Table 1 respectively.

3.2 Application of MNEDA for SRCSP

SRCSP can be defined as the following decision making process:

Operations sequencing: determine the sequence of executing all operations required for the jobs so that the precedence relationships among the all the operations are not violated.

Configuration selection: determine the machine

facility of operation which can be similarly processed on one of the multiple machine types candidates;

Resource Seizing: determine the accessory resources setup planning for operation according to the feature geometry and the available machining resources.

The decision of SRCSP can be mapped into three phases. Univariate marginal distribution is used to estimate the marginal probability, which represents the probability that job j will be scheduled at k-th locus of the operations sequence vector. This method is widely employed in recently research and development for stability improvement[17].

The second phase is to determine the machine configuration for each operation $o \in J_o$. As described in section 2, the decision of the set of o has explicit interdependence according to the related resource shared between the operations. Markov network is employed to estimates the interrelation of operations that share the same configuration.

The final phase is to seize the resource in order to reduce changeovers of resources among the machine configurations, Earliest Completion Time (ECT) based seizing rule-based heuristic algorithm is adopted to treat with resource loading and unloading, which enhanced computation by avoiding high-level complexity due to matching with resources and operations[18].

3.3 Results and Discussion

Optimality: The larger-scale problems were used to evaluate CEDA performance compared with current meta-heuristics, hybrid hrGA[6] and UMDA[17]. Parameters of simulation data were aligned in order to provide a comparative basis (shown in Table1). Fig. 4 shows box plots for the experimental result. It shows that MNEDA algorithms exhibited superior optimality to hrGA and

Table1. Parameters and strategies of hrGA UMDA and MNEDA

	hrGA	UMDA	MNEDA
Iteration	6000	6000	6000
Population	200	200	200
Selection	tournament(k)	-	-
Strategy	elitism	-	-
Operators	Crossover(P_m) Mutation(P_c)	Sampling Mutation(P_m)	Sampling
Parameters	$P_m = 0.30$ $P_c = 0.65$ $k = 2$	$\|D\| = 80\%$ learingRate= 0.02 $P_m = 0.30$	$\|D\| = 80\%$ $MN = 4$ $\alpha = 1.5$ $coolingRate = 0.5$

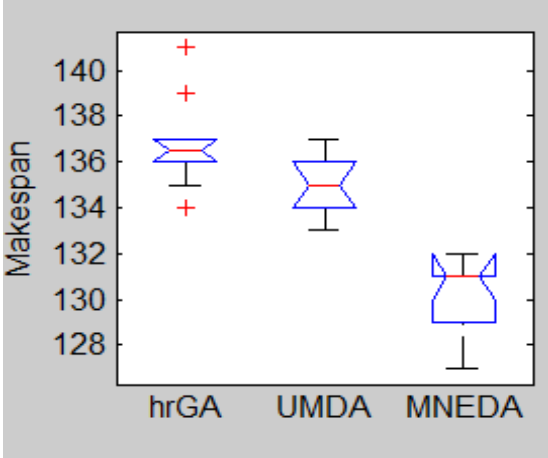


Fig4. Box plot for large-scale experiment

UMDA in experiment. The hrGA proposed the seizing-rules heuristic method, which considers the sequence-dependent changeover between machines. In contrast to the evolutionary operator of hrGA. UMDA and MNEDA uses the estimated probability distributions of decision variables relating to operation sequence and machine setup planning. It provides a prediction mechanism on the variant of decision variables. As shown in the figures, MNEDA models the interaction between the variables. MNEDA achieved superior stability than UMDA and hrGA.

Computation Cost: The computational costs of evolutionary-based algorithms mainly depend on the number of fitness evaluations. The difference of time complexity among hrGA, UMDA, and the proposed approach mainly relies on the operators. MNEDA samples a new candidate solution and improved the current candidate solution according to the probability distribution. Moreover, it estimates the structure of Markov network decision variables using the promising data. The complexity of estimating the structure in term of mutual information is $O(n^2)$.

The average CPU times of hrGA, UMDA and MNEDA were 870.86s, 1105.02s and 1245.17s, respectively. The results show that achieved although MNEDA similar computational efficiency to UMDA, it needed take some additional time estimating the structure of Markov network.

4. Conclusion

In this paper, we have proposed a local Markov property based EDA (MNEDA) for solving

sharable resource constrained scheduling problem. This avoids complex approximation in which involves the cliques and the corresponding potential functions. It can be implemented in simpler and easier Markov network based EDA. Considerable experiments were performed and the results confirmed that MNEDA outperforms recent research results for scheduling problems in FMS. Further research should be done in order to improve the performance of MNEDA, particularly on the implementation of more efficient structure learning algorithms, and also on the use of other more efficient versions of the Gibbs sampler algorithm.

References

[1] P. Brucker, A. Drexl, R. Möhring, K. Neumann, and E. Pesch, "Resource-constrained project scheduling: Notation, classification, models, and methods," *Eur. J. Oper. Res.*, vol. 112, no. 1, pp. 3–41, 1999.

[2] J. Böttcher, A. Drexl, R. Kolisch, and F. Salewski, "Project Scheduling Under Partially Renewable Resource Constraints," *Manag. Sci.*, vol. 45, no. 4, pp. 543–559, Apr. 1999.

[3] S. Hartmann and D. Briskorn, "A survey of variants and extensions of the resource-constrained project scheduling problem," *Eur. J. Oper. Res.*, vol. 207, no. 1, pp. 1–14, Nov. 2010.

[4] D. Gerwin, "Manufacturing flexibility: a strategic perspective," *Manag. Sci.*, vol. 39, no. 4, pp. 395–410, 1993.

[5] F. Ballestín and N. Trautmann, "An iterated-local-search heuristic for the resource-constrained weighted earliness-tardiness project scheduling problem," *Int. J. Prod. Res.*, vol. 46, no. 22, pp. 6231–6249, 2008.

[6] J.-Z. Wu and C.-F. Chien, "Modeling semiconductor testing job scheduling and dynamic testing machine configuration," *Expert Syst. Appl.*, vol. 35, no. 1–2, pp. 485–496, 2008.

[7] G. Singh and A. T. Ernst, "Resource constraint scheduling with a fractional shared resource," *Oper. Res. Lett.*, vol. 39, no. 5, pp. 363–368, Sep. 2011.

[8] S. Baluja and S. Davies, "Using optimal dependency-trees for combinatorial optimization: Learning the structure of the search space," in *Machine Learning International workshop Then Conference*, 1997, pp. 30–38.

[9] P. Larrañaga and J. A. Lozano, *Estimation of distribution algorithms: a new tool for evolutionary computation*. Springer, 2002.

[10] S. Shakya and J. McCall, "Optimization by estimation of distribution with DEUM framework based on Markov random fields," *Int. J. Autom. Comput.*, vol. 4, no. 3, pp. 262–272, Jul. 2007.

[11] M. Pelikan, "Bayesian Optimization Algorithm," *Hierarchical Bayesian Optim. Algorithm*, pp. 31–48, 2005.

[12] A. Freno, *Hybrid Random Fields*. Springer, 2011.

[13] X. Hao, H. W. Lin, X. Chen, and T. Murata, "Cooperative Bayesian Optimization Algorithm: a Novel Approach to Multiple Resources Scheduling Problem," *Ieej Trans. Electron. Inf. Syst.*, vol. 132, no. 12, pp. 2007–2018, 2012.

[14] S. Shakya, "DEUM: A framework for an Estimation of Distribution Algorithm based on Markov Random Fields," 2006.

[15] R. Santana, P. Larrañaga, and J. Lozano, "Research topics in discrete estimation of distribution algorithms based on factorizations," *Memetic Comput.*, vol. 1, no. 1, pp. 35–54, 2009.

[16] A. Allahverdi, C. T. Ng, T. C. E. Cheng, and M. Y. Kovalyov, "A survey of scheduling problems with setup times or costs," *Eur. J. Oper. Res.*, vol. 187, no. 3, pp. 985–1032, Jun. 2008.

[17] L. Wang, S. Wang, Y. Xu, G. Zhou, and M. Liu, "A bi-population based estimation of distribution algorithm for the flexible job-shop scheduling problem," *Comput. Ind. Eng.*, vol. 62, no. 4, pp. 917–926, May 2012.

[18] A. Allahverdi, J. N. . Gupta, and T. Aldowaisan, "A review of scheduling research involving setup considerations," *Omega*, vol. 27, no. 2, pp. 219–239, Apr. 1999.

New Best Solutions for the Close-Open Mixed Vehicle Routing Problem by using an Improved Harmony Search Algorithm

Tantikorn Pichpibul **Ruengsak Kawtummachai**

Faculty of Business Administration
Panyapiwat Institute of Management, Thailand
{pichpibul@gmail.com, ruengsak@yahoo.com}

Abstract

This paper addresses the close-open mixed vehicle routing problem (COMVRP) that both close and open variants of the vehicle routing problem are simultaneously considered. The objective is to find the solution which minimizes the total cost of delivery. In this paper, we propose an improved harmony search algorithm (IHSA) composed of two main approaches. First, the probabilistic Clarke-Wright savings algorithm is applied to improve an initial COMVRP solution. Second, that solution is used to produce the best COMVRP solution by the music improvisation process. The computational results on well-known benchmark problems show that the IHSA is superior to the best existing algorithm. Moreover, new best solutions for 10 benchmark instances are also found.

Keywords: Close-open mixed vehicle routing problem, Clarke-Wright savings algorithm, Harmony search algorithm

1 Introduction

The close-open mixed vehicle routing problem (COMVRP) [1] is a new variant of the vehicle routing problems which is the combination between the capacitated vehicle routing problem (CVRP) [2] and the open vehicle routing problem (OVRP) [3]. Both CVRP and OVRP are known as NP-hard [4] [5]. Therefore, COMVRP is also a NP-hard. Liu and Jiang [1] mentioned that COMVRP is more complex and even harder to be tackled than CVRP and OVRP. The COMVRP arises in the real-world applications. For example, the companies which have their own trucks deliver their products to many customers with fluctuant demand. But in the situation that the total demand of customers is greater than the total capacity of own trucks, the decision makers may consider to deal with third party carriers. In COMVRP, their own trucks must return to their warehouse after completing the delivery which is one characteristic of CVRP, but the outsourcing trucks do not return to the warehouse and stop the delivery at the last customer which is one characteristic of OVRP. The other characteristics of COMVRP can be described as the problem of determining a set of vehicle routes to serve a set of customers with known demand. Each route represents a sequence of customers that must be visited by a vehicle. The COMVRP involves a single warehouse, a homogeneous fleet of vehicles and a set of customers. The objective is to find the solution which minimizes the total cost of delivery while serving all customers' demand. The example routes for COMVRP are shown in Figure 1.

2 Literature Review

The vehicle routing problems have attracted many researchers since Dantzing and Ramser [2] proposed the problem in 1959. Many efficient heuristics were presented to solve CVRP and OVRP, such as a tabu search [6] [5], an ant colony system [7] [8], a particle swarm optimization [9] [10], and a genetic algorithm [11] [12]. Although CVRP and OVRP were tackled by many researchers, Liu and Jiang [1] is only one who tackled COMVRP by using a memetic algorithm. In the literature, we have found that a harmony search algorithm (HSA) proposed by Geem *et al.* [13] is one of recent algorithms which had successful applications to various optimization problems [14] [15]. However, there is no previously published research that presents HSA to solve the vehicle routing problems. This is our contribution to develop a new HSA which can efficiently solve the COMVRP in terms of solution quality.

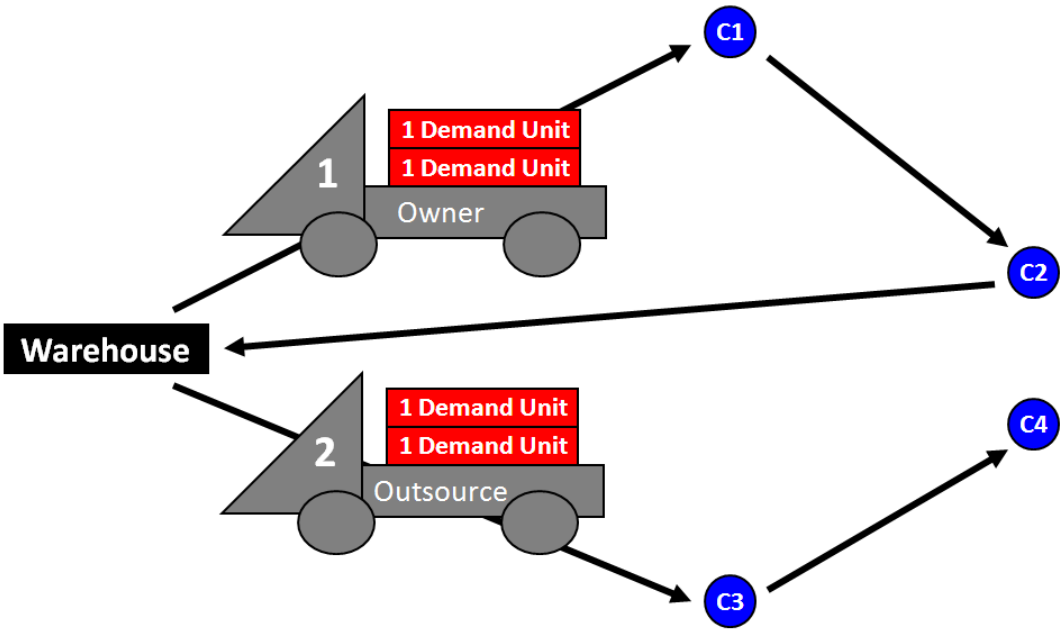


Figure 1. The example routes for COMVRP

3 Method of Approach

The development of our approach based on the music improvisation process in HSA [13] and the probabilistic Clarke-Wright savings algorithm [16] that the procedures are described in the following subsections.

3.1 The Harmony Search Algorithm

The harmony search algorithm (HSA) [13] is simulated by the process of music improvisation when a group of musicians playing musical instruments together. Each musician sounds any pitch within the possible range of notes. The composition of notes makes the harmony vectors which are kept in the harmony memory (HM). Each harmony vector is estimated by an aesthetic standard. In the improvisation process, a new harmony vector is made by using memory consideration, pitch adjustment and random selection. If the new harmony vector is better than the worst harmony vector in HM, it is updated to HM instead of the worst one. This process is repeated until a perfect harmony is found by the musicians.

The improvements of HSA in this paper were focused on the modification of parameters and procedures of HSA to deal with COMVRP. Moreover, we also improved the solution quality of COMVRP in HM. The details of our approach are described as follows.

The parameters are composed of range of notes which is a set of customers; harmony memory considering rate (HMCR) which is a probability number to choose a customer from the harmony vectors for the memory consideration; pitch adjusting rate (PAR) is a probability number to choose a neighboring customers from the customer of memory consideration.

The structures of the harmony vector according to the parameters were modified as shown in Figure 2. The routes in Figure 1 are presented as an example solution. We assume one warehouse and four customers. The number of notes in each harmony vector is also equal to four based on the number of customers. As shown in Figure 2, the total distance of the harmony vector for COMVRP is calculated by considering the composition of notes. In general, HSA always randomly generates the notes in each harmony vector. In contrast, we generate the notes in the first harmony vector by using a probabilistic Clarke-Wright savings algorithm proposed by Pichpibul and Kawtummachai [16].

In the improvisation process, the customer in the first note of all harmony vectors is chosen to be a new customer with equal probability by the memory consideration with a probability of HMCR. After the new customer is chosen from the memory consideration, it is shifted to be a new customer, which is neighboring customers composed of an upper customer and a lower customer with equal probability by the pitch ad-

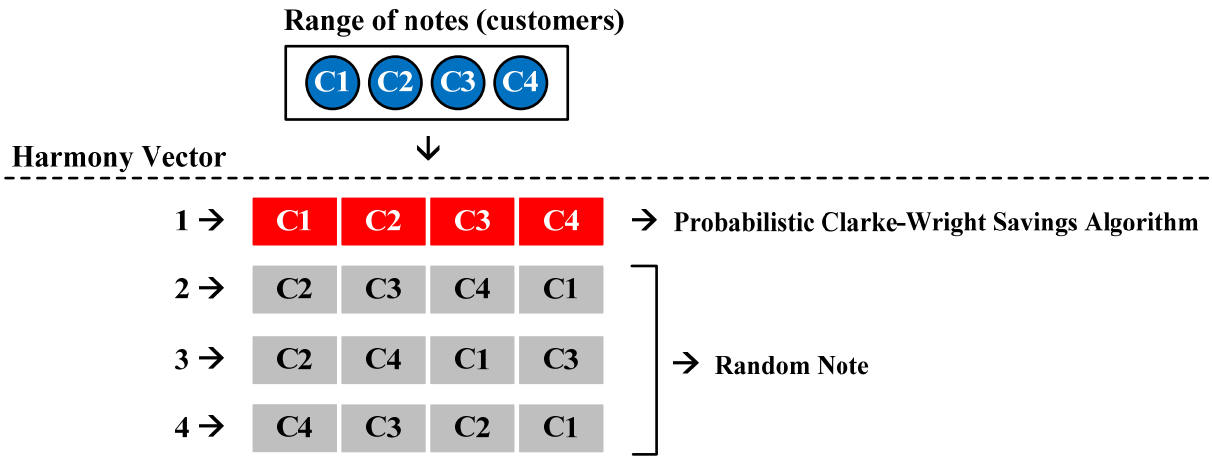


Figure 2. The structures of the harmony vector

justment with a probability of PAR. For the next note, the new customer is chosen within the possible range of customers by the random selection with a probability of 1-HMCR. This process is repeated to the next note in the harmony vectors until the last note is reached.

3.2 The Probabilistic Clarke-Wright Savings Algorithm

The Clarke-Wright savings algorithm first calculates the savings value for every pair of customers as follows:

$$s_{i,j} = c_{0,i} + c_{j,0} - c_{i,j} \tag{1}$$

Here, $c_{i,j}$ is the traveling distance between customers i and j and the warehouse is assigned to be customer 0. The savings is thus calculated and then ranked in the descending order to the savings list. Next, in each savings value from the list, any two customers i and j are combined to the route. This procedure is repeated to process the next savings value in the list until no feasible route is possible or the last savings value is reached. In this paper, instead of ranking the savings value in the descending order, we randomly rank the value in the savings list by two-phase probabilistic procedure which is an iterative improvement approach that rearranges the new cost savings value (more details are available in [16]).

4 Computational Results

The IHSA was coded in Visual Basic 6.0 on an Intel® Core™ i7 CPU 860 clocked at 2.80 GHz with 1.99 GB of RAM under Windows XP platform. The numerical experiment used well-known benchmark problems composed of 15 instances with different features, e.g., uniformly and not uniformly dispersed customers, clustered and not clustered, with a centered or not centered depot. The first 11 instances are from Augerat *et al.* [17] in data sets A, B, and P. The second 4 instances are from Christofides and Eilon [18] in data set E. Table 1 shows the characteristics of the benchmark problems of CVRP, OVRP, and COMVRP.

The numerical results are also shown in Table 2. Out of 15 instances, we find the new best solutions for 10 instances, which are highlighted by using bold type, and similar solutions for 5 instances. Because Liu and Jiang [1] did not report their best solutions in the paper, we therefore provide the example of our best solutions as shown in Table 3, and the routes are also illustrated in Figure 3.

5 Conclusions

In this paper, we have proposed an improved harmony search algorithm (IHSA) to solve the close-open mixed vehicle routing problem (COMVRP). We have modified IHSA with the music improvisation process and the probabilistic Clarke-Wright savings algorithm. We also have done the experiments by using well-known benchmark problems composed of 15 instances obtained from the literature, and have compared them with the best known solutions. The computational results show that IHSA is superior to the best existing heuristic in terms of solution quality. It generates the new best solutions for 10

Table 1. Characteristics of the benchmark problems of CVRP, OVRP, and COMVRP

Instances	N [a]	Q [b]	CVRP		OVRP		COMVRP		
			V_1 [c]	V_2 [d]	V_1 [c]	V_2 [d]	D [e]	V_1 [c]	V_2 [d]
A-n32-k5	32	100	5	0	0	5	246.89	3	5
A-n33-k5	33	100	5	0	0	5	176.68	3	5
A-n33-k6	33	100	6	0	0	6	175.83	3	6
A-n34-k5	34	100	5	0	0	5	201.73	3	5
A-n36-k5	36	100	5	0	0	5	236.53	3	5
B-n51-k7	51	100	7	0	0	7	186.69	4	7
B-n52-k7	52	100	7	0	0	7	158.17	4	7
B-n56-k7	56	100	7	0	0	7	186.22	4	7
B-n57-k7	57	100	7	0	0	7	197.65	4	7
P-n21-k2	21	160	2	0	0	2	132.10	1	2
P-n22-k8	22	3000	8	0	0	8	109.71	5	8
E-n22-k4	22	6000	4	0	0	4	110.47	2	4
E-n30-k3	30	4500	3	0	0	3	215.12	2	3
E-n33-k4	33	8000	4	0	0	4	262.48	2	4
E-n51-k5	51	160	5	0	0	5	135.77	3	5

[a] Number of customers, [b] Vehicle capacity, [c] Number of available own vehicles, [d] Number of available outsourcing vehicles, [e] Traveling cost constraint

Table 2. Computational results

Instances	CVRP			OVRP			COMVRP		
	OPT [19]	IHSA	%Dev[a]	OPT [19]	IHSA	%Dev[a]	BKS [1]	IHSA	%Dev[b]
A-n32-k5	787.08[a]	787.08	0.000	487.31[a]	487.31	0.000	1063.53[b]	1063.53	0.000
A-n33-k5	662.11[a]	662.11	0.000	424.54[a]	424.54	0.000	1004.49[b]	**994.91**	**-0.954**
A-n33-k6	742.69[a]	742.69	0.000	462.43[a]	462.43	0.000	1170.88[b]	**1169.29**	**-0.136**
A-n34-k5	780.94[a]	780.94	0.000	508.17[a]	508.17	0.000	1102.98[b]	**1102.17**	**-0.074**
A-n36-k5	802.13[a]	802.13	0.000	519.46[a]	519.46	0.000	1096.35[b]	1096.35	0.000
B-n51-k7	1035.04[a]	1035.04	0.000	625.14[a]	625.14	0.000	1494.54[b]	1494.54	0.000
B-n52-k7	749.97[a]	749.97	0.000	441.19[a]	441.19	0.000	1208.93[b]	**1202.70**	**-0.515**
B-n56-k7	712.92[a]	712.92	0.000	420.49[a]	420.49	0.000	1169.84[b]	**1168.22**	**-0.138**
B-n57-k7	1157.73[a]	1157.73	0.000	646.36[a]	646.36	0.000	1620.51[b]	**1549.48**	**-4.383**
P-n21-k2	212.71[a]	212.71	0.000	163.88[a]	163.88	0.000	383.46[b]	383.46	0.000
P-n22-k8	600.83[a]	600.83	0.000	345.87[a]	345.87	0.000	1183.87[b]	1183.87	0.000
E-n22-k4	375.28[a]	375.28	0.000	252.61[a]	252.61	0.000	700.94[b]	**700.93**	**-0.001**
E-n30-k3	535.80[a]	535.80	0.000	393.51[a]	393.51	0.000	733.24[b]	**730.83**	**-0.329**
E-n33-k4	837.67[a]	837.67	0.000	511.26[a]	511.26	0.000	1065.46[b]	**1057.05**	**-0.789**
E-n51-k5	524.61[a]	524.61	0.000	416.06[a]	416.06	0.000	929.70[b]	**921.28**	**-0.906**

[a] Percentage deviation between the optimal solution (OPT) and IHSA
[b] Percentage deviation between the best known solution (BKS) and IHSA

Table 3. New best solution for A-n33-k5 instance

No	Type	Route	Cost
1	CVRP	0-2-32-13-8-7-26-5-12-0	222.660
2	CVRP	0-11-31-1-21-14-19-6-24-0	169.301
3	CVRP	0-23-18-28-22-0	96.623
4	OVRP	0-20-4-10-30-25-27	232.952
5	OVRP	0-15-29-16-3-9-17	273.372
		Total cost	994.908

instances and makes tie with the existing best known solutions for 5 instances. Therefore, we can conclude that the performance of IHSA is excellent for solving COMVRP.

References

[1] R. Liu, and Z. Jiang. The close-open mixed vehicle routing problem. *European Journal of Operational Research*, 220:349-360, 2012.

[2] G.B. Dantzig, and J.H. Ramser. The truck dispatching problem. *Management Science*, 6:80-91, 1959.

[3] D. Sariklis, and S. Powell. A Heuristic Method for the Open Vehicle Routing Problem. *The Journal of the Operational Research Society*, 51:564-573, 2000.

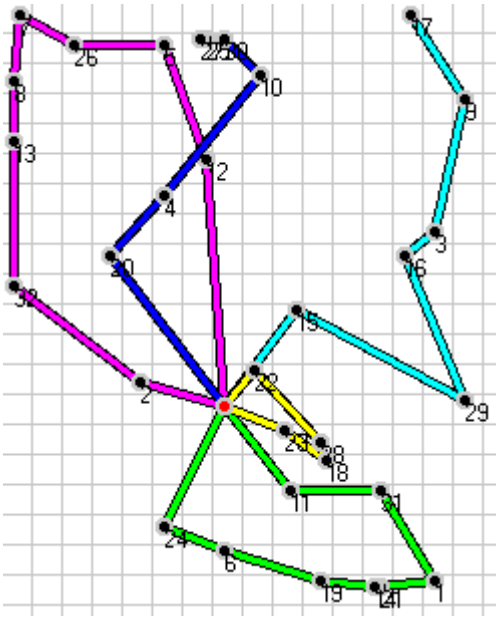


Figure 3. The routes for A-n33-k5 instance

[4] P. Toth, and D. Vigo. *The vehicle routing problem*. SIAM Publishing, Philadelphia, PA, 2002.

[5] J. Brandão. A tabu search algorithm for the open vehicle routing problem. *European Journal of Operational Research*, 157:552-564, 2004.

[6] Y. Rochat, and É.D. Taillard. Probabilistic diversification and intensification in local search for vehicle routing. *Journal of Heuristics*, 1:147-167, 1995.

[7] M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Mans, and Cybernetics*, 1:29-41, 1996.

[8] X. Li, and P. Tian. *An ant colony system for the open vehicle routing problem*. In: Dorigo et al., editor. Lecture Notes in Computer Science (ANTS 2006) 4150, Berlin: Springer, 356-63, 2006.

[9] T.J. Ai, and V. Kachitvichyanukul. Particle swarm optimization and two solution representations for solving the capacitated vehicle routing problem. *Computers & Industrial Engineering*, 56:380-387, 2009.

[10] S.A. MirHassani, and N. Abolghasemi. A particle swarm optimization algorithm for open vehicle routing problem. *Expert Systems with Applications*, 38:11547-11551, 2011.

[11] M. Gen, and R. Cheng. *Genetic algorithms and engineering design*. Wiley, New York, 1997.

[12] S. Yu, C. Ding, and K. Zhu. A hybrid GA-TS algorithm for open vehicle routing optimization of coal mines material. *Expert Systems with Applications*, 38:10568-10573, 2011.

[13] Z.W. Geem, J.H. Kim, and G.V. Loganathan. A New Heuristic Optimization Algorithm: Harmony Search. *Simulation*, 76:60-68, 2001.

[14] K.S. Lee, and Z.W. Geem. A New Meta-Heuristic Algorithm for Continuous Engineering Optimization: Harmony Search Theory and Practice. *Computer Methods in Applied Mechanics and Engineering*, 194:3902-3933, 2005.

[15] Z.W. Geem, M. Fesanghary, J. Choi, M.P. Saka, J.C. Williams, M.T. Ayvaz, L. Li, S. Ryu, and A. Vasebi. *Recent Advances in Harmony Search*, in Advances in Evolutionary Algorithms, W. Kosinski, Ed. Vienna: I-Tech Education and Publishing, 2008.

[16] T. Pichpibul, and R. Kawtummachai. An improved Clarke and Wright savings algorithm for the capacitated vehicle routing problem. *ScienceAsia*, 38:307-318, 2012.

[17] P. Augerat, J. Belenguer, E. Benavent, A. Corbern, D. Naddef, and G. Rinaldi. *Computational results with a branch and cut code for the capacitated vehicle routing problem*. Grenoble, France: Universite Joseph Fourier, 1995.

[18] N. Christofides, and S. Eilon. An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, 20:309-318, 1969.

[19] A.N. Letchford, J. Lysgaard, and R.W. Eglese. A branch-and-cut algorithm for the capacitated open vehicle routing problem. *Journal of the Operational Research Society*, 58:1642-1651, 2007.

Cognitive technologies for assessing skills

Charles Tijus † François Jouen † Sébastien Poitrenaud † Michèle Molina ‡

†Cognitions Humaine et Artificielle (CHArt), EPHE et Université Paris 8, Paris, France
‡ PALM, Université de Caen Basse-Normandie, Caen, France
{tijus, poitrenaud}@univ-paris8.fr, Francois.Jouen@ephe.sorbonne.fr,
michele.molina@unicaen.fr

Abstract

Solving any task involves thinking-to-gestures processes: unobservable mental superordinate goals to primitive hand movements acting on objects as trajectory data. The question is how to evaluate know-how taking into account both decision-making and dexterity. From the Cognitive Science literature and considering the collection of relevant data, we derive a functional definition of Know-How that can be useful for the conception of an automatic assessment system of Skills acquisition. This functional definition is also used to introduce a computational cognitive model that is based on the problem solving and task analysis literature. Discussion is about how much the model is of help for the building of a platform to implement the method as part of a MOOC.

Keywords: Model of Cognition, Task Analysis, Problem-solving, Complex Systems, eLearning, Cognitive technologies, Massive Online Open Courses

1 Introduction

Since 2007, a new trend in eLearning is the merge of free online classes, the Massive Online Open Courses (MOOC).

MOOCs that are about declarative knowledge can provide ways to evaluate how much students increase their knowledge about facts and concepts. Optimizing learning can be done through adapting swarm intelligence systems into what was called the man-hill optimization paradigm [1][2][3]. This means that finding optimal paths through the graphs of educational contents, valued by the preceding students' successes and failures helps evaluating a particular student' path till now as well as advising about the next steps of the learning processes.

Automatic assessment of knowledge can also be done using evidential Multiple choice questions, - which is easy-, but also through assessment of free texts provided by students. Generic tools are Latent Semantic Analysis (LSA) and SUMMA-LSA, two models that maps meaning into a semantic space that can be used to compute the semantic distances between words, sentences, and whole texts [2][3].

For automatic assessment of free texts provided by students, beyond LSA, SUMMA-LSA, there are Topics Model [7] that use a hidden-Markov model to learn sequential dependencies to be combined with the semantic information from the topics representation [8]. The Topics Model allows computing how much a given free response is close to the one provided by the question provider, either a teacher or another tutor student.

The content of eLearning classes are about verbal, written, graphical information, named here declarative knowledge, but also about Know-how with demonstration of the "how-to-it", named here procedural knowledge.

The question under investigation is how we can model Know-how for eLearning. As a difference with declarative knowledge, Know-how can difficultly be captured with talks (typing): the procedural knowledge of Know-How is related to the behavior of an agent that uses objects, mainly in the context of learning for professional purposes.

The topic of the paper is about how to model distant Know-How in order to be captured for assessment.

2 What is Know-How

2.1 Know-How as Goal States

Know-how is the skill of an agent to perform tasks. Performing a task is to achieve a goal state

(the task is completed) [9].

- Example 1: Mail a letter => the letter is posted
- Example 2: Solve the Tower of Hanoi => the problem is resolved

This definition is that the result of a task is a "*state of the world*" in which the objects have the transformed properties defined by the goal.

Figure 1. Solving a task is a State of objects having their properties changed: the place of the 3 Disks of the Tower of Hanoi will be changed (from left to right).

Definition 1 – Know-How is an internal (mental) representation of the goal state of related task.

2.2 Know-How as sub-Goals Structure of Tasks

In order to fulfill a goal state of a task, the agent has to make actions to transform objects in a sequential temporal order.

Knowing the goal but not the procedure is problem solving. Knowing the goal and the set of sequential actions (named here action primitives) is running task. However, knowing the "*how-to-do-it*" for a task does not entail knowing the "*why-to-do-it*". The procedure has a sub-goals structure that provides its semantics. For instance, while the sequence of actions for the Tower Of Hanoi is as follows,

Example 2 - developed:

1. Move small disk to left place
2. Move medium disk to middle place
3. Move small disk to middle place
4. Move large disk to left place
5. Move small disc to right place
6. Move medium disc to left place
7. Move small disc to left place

The sub-goals structure is as follows.

1. Move large disk to left place
To do so: move medium and small disk to middle
    1.1. Move medium disk to middle place
        To do so move small disk to left
            1.1.1. Move small disk to left
    1.2. Move small disk to middle place
2. Move medium disc to left place
To do so: move small disk to right place
    2.1. Move small disc to right place
3. Move small disc to left place

Note that the tree-like sub-goals structure provides the semantic of the procedure: the how by the top-down path (*e.g., how to move the medium disk to the middle place: 1.1? By moving first the small disk to the left: 1.1.1*) and the why by the bottom-up path (*e.g., why do you move the medium disk to the middle: 1.1? To move the large disk to the left: 1, after having also moved the small disk to the middle place: 1.2*).

Traditionally tasks are modeled using such a tree-like sub-goals structure [10], [11] with logical (*IF, OR, XOR…*) and temporal (*No order, Same time, Until...*) operators (*or constructors*) that indicate time relationships between sub-goals. The set of sub-goals has a hierarchical structure: high levels sub-goals are found at the top of the tree while the observable action units (*primitive*) are at the leaf nodes.

Definition 2 – Know-How is about how to perform any particular sub-goal as well as being able to explain why this sub-goal is to be performed.

2.3 Know-How as a set of conditions, prerequisites, constraints and post-requisites

Know-how is not solely being able to know what to do (*goal*), how to do it (*procedure as sequences of actions on objects*) and why (*sub-goal structures*). Know-how is also the taking into account of external constraints (*Do I get sufficient time for performing that task? Is it the priority among tasks?*) as well as internal constraints (*I have to stay in that room*).

More crucial are conditions (*things that need to be true otherwise the task cannot be performed: e.g., the necessity of having a specific tool*), prerequisites (*to be done before starting running the task*) and post-requisites (*to be done after the task is achieved*).

Consider the following list of ordered actions, provided as professional procedure, to perform a medical test on patients for checking lack of insulin.

Example 3:

- Wash hands before and after treatment. It is a non-sterile treatment.
- Explain to the patient the course and purpose of the treatment.
- Have the patient urinate directly into the goblet to the extent possible, otherwise use a urinal or basin.
- Put non-sterile gloves.
- Dip the strip into the urine.
- Wait the recommended time for reading the test strip (*usually a few tens of seconds*).
- Compare the resulting color with the color scale on the vial.
- Record the result of the care record and send it to the nurse.
- Dispose of waste, cleaning and disinfecting equipment. Precautions
- Check the expiry date of the test strip vial.
- Do not touch the test part of the strip.
- Observe the reading time.

The understanding of this task requires differentiating

- CONDITIONS: having gloves, test strips with correct expiry date, a goblet, urinal or basin,
- PREREQUISITES: explain to the patient, have the patient urinate, Wash hands, Put sterile gloves,
- INTERNAL CONSTRAINTS: do not touch the test part of the strip, observe the reading time,
- POST-REQUISITES: Dispose of waste, cleaning and disinfecting equipment, wash hands.

The GOAL is "having the recorded result sent to the nurse" and the sub-goals are: 1 - dip the strip into the urine, 2 - wait the recommended time for reading the test strip, 3 - compare the resulting color with the color scale on the vial, 4 - record the result of the care record and 5 - send it to the nurse. Thus, From the GOAL and from the sub-goals the PROCEDURE is to be built in a tree-like goal structure, as follows:

Example 3 - revisited:

1. Make the test strip reacts
    1.1 - dip the strip into the urine,
    1.2 - wait the recommended time for reading the test strip
2. Record value of strip reaction
    2.1 - compare the resulting color with the color scale
    2.2 - record the result of the care record
3 - send record to the nurse.

Definition 3 – Know-How is about the context of the task that allows performing it.

2.4 Know-How as a causal agent-to-patient objects chain

We have seen that the goal of a task is to have object(s) transformed. These goal objects are patient target objects that must be transformed for the purpose of the task: eg "*the letter must be mailed*", "*the three pieces of the tower of Hanoi to be right*", "*the result of the test is to be sent to the nurse*". This patient target object is transformed because of action made by functional objects.

Although the hand is usually the main functional object to act on patient objects, - the simplest case is its direct action on the patient target object: "*to take a knife*"-, the procedure of a task (*structured goals*) is to act on patient objects with functional objects as tools, or means that are used to obtain the final goal state. These functional objects (F) operate on patient objects (P) in a causal chain, where each object is in turn patient then functional [12], as follows:

Example 4:

F: HAND -> P: KNIFE (handle)
F: KNIFE (Blade) -> P: BREAD (outside)
F: BREAD (container)-> P: CHEESE (content)
=> SANDWICH

All along the running task, objects are used for their function to operate on other patient objects. First, there are relations between functional and patient objects: a knife is used to cut the meat because the meat is sufficiently soft to be cut. Thus, performing a task is to construct a causal chain involving patient objects that are transformed to be used as functional objects: e.g., the hand (F) takes the knife (P) in order to use the knife (F) to cut the meat (P).

Second, an object can be used in turn as patient then functional object because it has both patient and functional parts (Fig. 2).

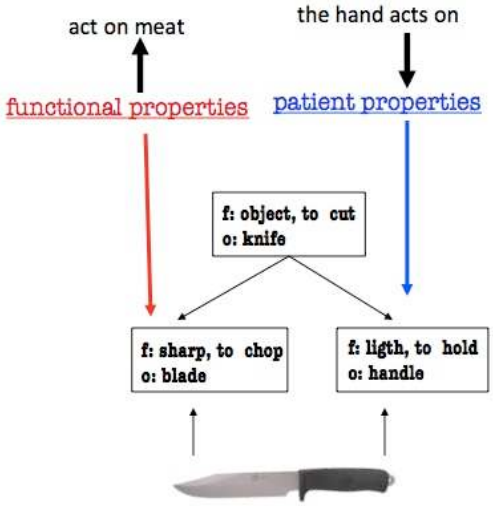


Figure 2. The whole-t-part decomposition of Tasks objects: parts are either patient (*the handle of the knife*) or functional (*the blade of the Knife*).

Definition 4 – Know-How is the knowledge of the causal chain involving patient and functional parts of each object as well as the relation between the functional part of a F object and the patient parts of a P object on which the former is operating.

2.5 Know-How as a causal agent-to-patient objects chain

As a matter of fact, the know-how about the preceding task medical test example is not delivered alone. It is delivered with some knowledge as follows:

Example 3 - completed:
"When the food shortage, the body will draw on its reserves (in the liver and muscles). After exhaustion of these reserves, there will be use of fats and proteins, whose degradation produces ketones. For diabetic patients with a lack of insulin (*which allows the passage of glucose into the cells*), the cells of the body, lacking of glucose, will use the fat reserves of where the production of ketones. These are hazardous waste for the body that urine eliminates". The professional who is performing the medical test and has this knowledge should be able to link know-how with this knowledge, being able to know what the test trip is measuring.

Definition 5 – Know-How is the competency to link the procedure with the domain knowledge of the task.

3 Know-how data for eLearning Assessment

A proper way for a teacher to evaluate a student know-how acquisition is to observe her realizing tasks and to ask questions about the understanding of how and why. Distant teaching of many students does not make it possible as well as the making of videos by students that may not have the required tools.
Although incomplete, because virtual objects do not have all of the properties of physical ones, Serious Games (SG) can be used, in addition to eLearning classes, for practicing, acquiring and assessing Know-How.
The specific question under investigation is how we can model the behavior of students for Know-how acquisition through for eLearning. The behavior of distant students is a number of dependent variables that we record for the Know-how Digital Campus (KHDC) project.

3.1 Know-How data as Human-Machine interaction

Learners have to display to themselves eLearning contents using motor commands (touching, clicking…).
We developed SiouXtrack with a traces recording module that allows the observation of all user actions. This module takes the form of an extension to the web browser installed on the client device. All interactions between the user and the browser are recordable in the form of user actions and subsequent responses of the system with more or less detail. Any action on the interface of the viewed website can be compared with the system events triggered and the specific consequences that follow for the user. Usual widgets like buttons, lists, tables and menus, text inputs and multimedia items like sounds, images and videos are taken into account. Context switches like opening and closing of documents, changes and manipulations of windows and tabs are also recorded.

Any information needed for offline simulation of the user experience can be recorded and stored into the SiouXtrack database. An interface for the observer-researcher accessed from the browser preference pane allows entering of parameters to fine tune the way traces are to be collected: list of events to consider, sharpness of recordings, properties of the web page elements to be stored and more.
Viewed URLs filtering can be turned on, using regular expressions, to define precisely the set of URLs for which user behavior is to be recorded.
The collected traces are transmitted to a server and stored in the SiouXtrack database. This database can then be queried to provide aggregated statistical analyzes and summaries of types of user paths as well as analysis in terms of individual protocols.

3.2 Know-How data as Eye-movements

Once contents are displayed, learners have to read and watch. We also developed eye-tracking technology from the webcam. What was read, seen, inspected and for how long is under the scope of the subsequent online analysis that can be done from eyes movements recording [13].
Eye-tracking method [*see 14 for a large review*] is the eye-movement recording when a task. Generally, a key task activity is visual search. Eye-tracking analysis allow exploring visual search of a particular information through exploration and the processing of the objects on the interface. The exploration corresponds to the transition between different areas of the page of the site, such as navigation menu, text, www links, picture, video…
Thus, eye-tracking allows to identify the areas seeing by subjects, the areas never explored by them, the area first explored, etc. [15]. The processing of eLearning-objects leads to search for particular information to decide if related to the main goal or to the current sub-goal.
With this method, the scan path, - that is the fixation-saccade succession, with progressive (*towards unexplored area*) or regressive (*towards an area previously fixated*) saccades -, allows knowing where the student looks and for how long. Thus, the completeness of the collection of information, but also the order in which the information is collected can be evaluated regarding the procedure of the running task to be watched or assessed.

3.3 Know-How data as body movement and hand gestures

Know-how is also the primitive real actions that have to be performed. The correct sequence, velocity, time delays, precision are variables of the body movements and of hand gestures. For this Human behavior recognition, we use 3D Zernike Moments approach [16, Fig. 3].
Once the body or hands movements in videos are recognized, correctness, velocity (speed and acceleration), time delays and precision can me measured by comparing the path in 3D space with the one of the expert of the task.
This can be done if the students are wearing Inertial Measurement Units (IMUs), - a device that combines accelerometers and gyroscopes -, that transmit velocity as well as orientation in 3D space (fig. 4).

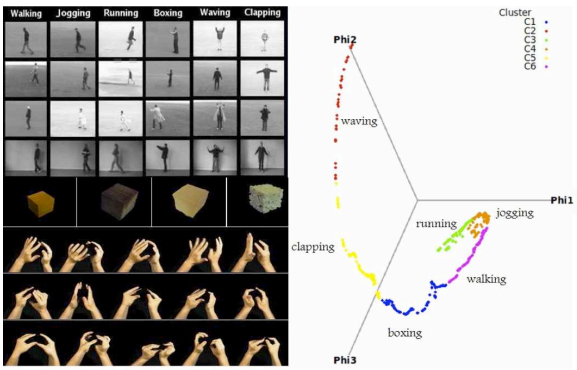


Figure 3. Classification of body and of hands movements, using a 3D Zernike Moments approach (*Bouziane, Chahir, Molina & Jouen, 2012*).

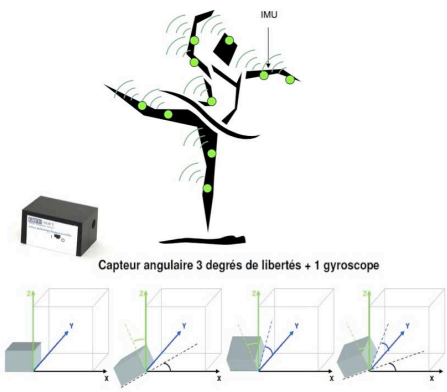


Figure 4. IMUs located on the body, harms and hands to record the spatial paths of gestures.

Thus, the spatial path done by students, when practicing, or for assessment, can be matched with the spatial path done by the professional of the task (fig. 5).

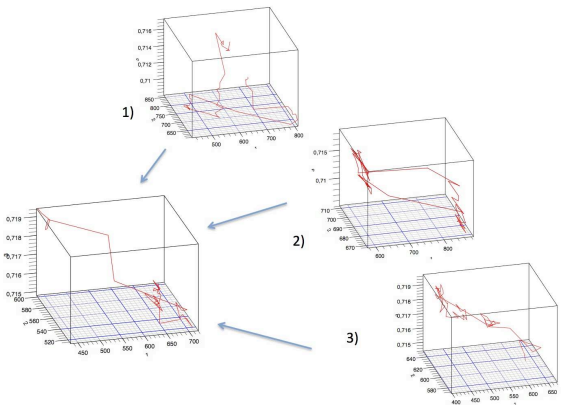


Figure 5. The recording of absolute 3D spatial trajectories done by a professional (*on left*) and done by a student (*on right*: three successive trials).

Measurement of similarity among different 3D spatial trajectories can be simply achieved by defining a spatiotemporal corridor and by deciding if a given trajectory is mainly within the corridor, and if late or in advance [17].
A more sophisticated technique is, - contrary to [18] clustering method -: (i) the segmentation of the pathway of the trajectory in basic predefined segment units. A pathway can then be defined as the ordered segment units set that composes it; each segment having a duration; (2) the use of Support Vector for calculate similarity; a trajectory being a vector whose dimension is the number of distinct segment units in it.
Thus, advising a student how to acquire the right gesture when acting is to check if her trajectory is in the expert corridor and, for precision, how much similar is the trajectory to the expert one.

4 Modeling Know-How for assessment

According to the definitions of what "Know-How" is (*section 2 above*) and the dependent variables that we can record (*section 3 above*), the remaining question is how to assess the Know-How that was acquired by an individual? The model we propose for its automatic assessment is hierarchical and modular.
First module is [*Goal Knowledge*]: the properties that the target task object must have in order for the task has been done. This module is independent because you can have a correct representation of content without knowing how to do to get it. This property of the task goal makes possible to evaluate if the task has been done, to use different procedures and materials, to make it done by another agent, etc. This module is also conditioning the evaluation of Know-How: why someone would be tested about a task if s/he does not know for what? This task goal level is known as being independent of procedure and of devices [19] [20].
Second module is the [*set of independent sub-goals*] (or subtasks). Tasks are made of sub-tasks that can be separately known without knowing how to link them to perform the task. For instance, the Tower of Hanoi problem is not about how to move each disk but in what order. This is an economical property for MOOCs (made of separate bricks of content) and for assessment of Know-How.
Third module is the knowledge [*task objects*], in terms of their patient and functional properties [12]. One might know how each device tool is functioning and how to use each device separately of others without knowing how to apply each tool to others.
Fourth module is the cognitive and goal-oriented Know-How module that is the built-in [*mental model]* of the task, made of the sub-goals structure, linking sub-goals to objects (not the reverse according to the object primacy [21]). This mental model, of a schema form, entails heuristics, inferences and planning by both top-down (the task goal) and bottom-up processes (perception of task objects) by simultaneously establishing the causal links between objects and the sub-goals structure.
Fifth module is [*perception*], visual, auditory, tactile, proprioception as inputs for capturing object information as well as feedback of how the running task is transforming objects.
Sixth module is [*body movements and gesture*], directed to objects, as first level output. Second output level is the state transformation of objects.
Last but not least, the seventh module is the [*metacognitive control of behavior*] dealing with external constraints, conditions, pre-requisites and post-requisites, according to the task specific goal, but also to more super-ordinate goals.

Figure 6 shows how these modules interact. Finally, for a given task, each combination of modules content can be matched with the domain knowledge in Long Term Memory (not displayed in the model depicted in Fig. 6, that concerns Working Memory).

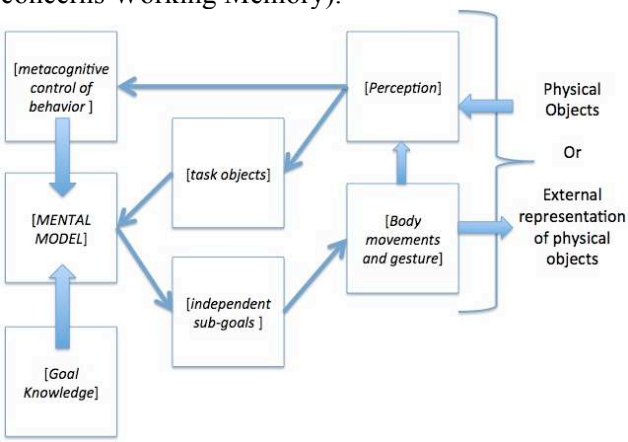


Figure 6. The modular model of Know-how: how the seven modules interact.

Although the model is a learning model, it prefigures the automaticity of the task with experience for which [*Perception*] and [*Body Movements and Gesture*] are more and more associated in such a way they are less and less associated to the other modules. The model also prefigures affordances, when the perception of an object indicates directly which of the sub-goals is to be used, avoiding decision-making in the mental model module.

5 Discussion and conclusion: Assessing Know-How

The modular model of Know-How provides the method for automatic assessment of the acquisition of running tasks.

First, before assessing a particular task, check if the task objects and the independent sub-goals are known. It happens that they are learnt while practicing the whole task. It is of interest to assess their learning per se. The assessment process is the same than for the whole task (next steps).

Second, verify the Goal Knowledge is known in order to avoid misunderstanding of what is to be done. Because you cannot pursuit the assessment, decide between providing the correct task goal or learning more.

Third, if the Goal task is known, question the How and why of the sub-goal structure. This can be done by questionnaire with multiple choices or by asking production of free texts.

Fourth, ask to perform the task in the virtual world of a serious game while recording eye movements [*Perception*] and actimetry [*body movement and gesture*]. Categorize the trajectories and evaluate how much their relative values fit the professional ones, either by the corridor technique or by computing vector similarity.

Fifth, correlate perception (*eye tracking*) and primitive-actions results (*body movements and gesture*) with the how and why of cognitive questions, looking for discrepancies.

Sixth, consider meta-cognitive control of behavior as individual parameters that rely either on personality or on the involvement in the task (motivation, emotion).

Is learning professional skills under the scope of eLearning and of MOOCs? Contrary to declarative knowledge that is mediated by symbolic external representation (words, graphs, diagrams, pictograms, images), Know-How is the action on physical objects, among them humans.

Because a science of simulation is still missing, the one that develops theories able to provide criteria about how much a simulation of X is a sufficient substitute of X, eLearning skills, obtained by using virtual objects in serious games, cannot ascertain the professional skills in the physical world. However, it may help facilitating the acquisition of Know-How for next real practices.

The Know-how Digital Campus (KHDC), -which is a MOOC that we are putting in place for learning many kinds of skills-, needs assessment techniques.

By considering what Know-How is and the collection of relevant data, we provided a model issued from the problem-solving and task-analysis literature. It is of help for implementing the whole assessment method in the next platform we are building.

References

[1] G. Valigiani, Y. Jamont, R. Biojout, E. Lutton and P. Collet. Experimenting with a Real-Size Man-Hill to Optimize Pedagogical Paths. In H. Haddad et al., Eds., *Symposium on Applied Computing, ACM*, Santa Fe, New Mexico, 2005

[2] G. Valigiani, E. Lutton and P. Collet. Adapting the ELO Rating System to Competing Subpopulations in a Man-Hill. *Proceedings of the 13th ISPE International Conference on Concurrent Engineering*, 2006, Antibes

[3] W. Lung-Hsiang and L. Chee-Kit. Swarm intelligence: new techniques for adaptive systems to provide learning support, *Interactive Learning Environments, 20:1*, 19-40, 2012

[4] T. K. Landauer and S. Dumais, S. A solution to plato's problem : The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological review, 104* :211-240, 1997

[5] T. K. Landauer, LSA as a Theory of Meaning. In T. K. Landauer, D. McNamara, S. Dennis, W. Kintsch (Eds). *The Handbook of Latent Semantic Analysis*, Lawrence Erlbaum Associates, Mahwah, 2007.

[6] W. Lenhard; H. Baier; J. Hoffmann, W. Schneider and A. Lenhard, A. Training of summarization skills via the use of content based feedback. In F. Wild, M. Kalz, J. van Bruggen & K. Kopper (Eds), *Proceedings of the First European Workshop on Latent Semantic Analysis Enhanced Learning,* pp.26-27, 2007. Heerlen : Open University of Netherlands.

[7] S. Jhean-Larose, V. Leclercq, J. Diaz, G. Denhière and B. Bouchon-Meunier (2010). Knowledge evaluation based on LSA : MCQs and free answers. *Studia Informatica Universalis, Mobile Ad-hoc NETworkS (MANETS),* Vol. 8.1., 53-78.

[8] T. L. Griffiths, M. Steyvers and J.B. Tenenbaum, J. B. Topics in Semantic Representation. *Psychological Review, 114(2),* 211-244, 2007.

[9] A. Zanga, J.F. Richard and Tijus, C. Implicit Learning in rule induction and problem solving. *Thinking & Reasoning, 10,* 55-83, 2004.

[10] S. Poitrenaud, S. The Procope Semantic Network: an alternative to action grammars. *International. Journal of Human-Computer Studies, 42*, 31-69, 1995

[11] S. Sebillotte. Methodology guide to task analysis with the goal of extracting relevant characteristics for human-computer interfaces. *International Journal of Human-Computer Interaction 7(4),* 341–363, 1995.

[12] E. Zibetti and C. Tijus. Understanding Actions: Contextual Dimensions and Heuristics. *Lectures Notes in Artificial Intelligenec vol. 3554, Modelling and Using Context.* New-York: Springer, 542-555, 2005.

[13] B.L. Nguyen, Y. Chahir , M. Molina and F. Jouen.Suivi du regard non intrusif en tête libre à partir d'images vidéo,, *Studia Informatica Universalis, Volume 8, N° 4,* pp. 31-57, 2010.

[14] K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124(3),* 372-422, 1998.

[15] Brézillon, P., Léger, L. and Tijus, C. (2009). Modeling users' search through contextual graphs. *Revue d'Intelligence Artificielle, 23*, 467- 484.

[16] A. Bouziane, Y. Chahir, M. Molina and F. Jouen. Unified Framework for Human behaviour recognition: An Approach using 3D Zernike moments, *Neurocomputing journal*, Elsevier, 100, 107-116, 2012.

[17] M. K. Essamgani. *Conception d'un outil de collecte de données sur les trajectoires*. Mémoire de Master 2. Document interne CHArt, 2013.

[18] M. K. El Mahrsi and F. Rossi. Graph-Based Approaches to Clustering Network-Constrained Trajectory Data In Proceedings. *Proceedings of the Workshop on New Frontiers in Mining Complex Patterns*, 2012), 184-195, Bristol, United Kingdom.

[19] J.F. Richard. *Logique du fonctionnement et logique de l'utilisation*. Rapport de recherche INRIA, N° 202, 1983.

[20] D.A. Norman, (1983). Some observations on mental models In D. Gentner and A.L. Stevens, (Eds), *Mental Models.* Hillsdale N.J.: Lawrence Erlbaum Associates, 1983, pp. 7–14.

[21] S. Poitrenaud, J.-F. Richard, and C. Tijus, Properties, categories and categorization. *Thinking and Reasoning, 11*, 151-208, 2005

Planning Support in Capability Growth Trajectory Mining

Thongchai Kaewkiriya*[1]*[2] Ryosuke Saga[2] Hiroshi Tsuji**[2]**

Faculty of Information Technology
Thai-Nichi Institute of Technology, Thailand*
Graduate School of Engineering
Osaka Prefecture University, Japan**
{thongchai}@tni.ac.th
{saga, tsuji}@cs.osakafu-u.ac.jp

Abstract

This study proposes two functions based on CMMI principles to guide a student's academic growth development plan by evaluating their growth history. One is "finding message generator" (FMG) function for dependency capability pairs and the other is "individuals plan generator" (IPG) function, which is based on a PDS (Plan–Do–See) cycle. The possibility of the proposed planning support functions is also discussed through flowchart diagrams, and an example based on 60 student's samples is provided.

Keywords: Trajectory mining, Data mining, planning support function.

1 Introduction

Data mining is becoming very popular in many areas, such as the analysis of business' growth [1], and is also being applied with personal e-learning systems to assist in online learning and teaching [2]. The consideration of the growth prospects for student's growth capability based on trajectory mining is an example of earlier research in which a student capability growth structure is extracted from growth trajectories [3]. The objective is to find growth patterns within student sample groups.

Capability structure extraction from growth trajectory to identify constraints on growth and to extract the capability structure of a student dataset has been proposed [4]. In the past, we proposed the formal concept of trajectory mining in the capability space to implement an information system [5]. A practical application involving software outsourcing demonstrated the proposed concept [1]. It used the basic concept of Capability Maturity Model Integration (CMMI) to evaluate software development on the basis of its capability and maturity [6]. Consideration of maturity and capability expands from initial development to optimization. Capability assessment involves five levels: Level 1 (the lowest) to Level 5 (the highest). In [7], the concept of CMMI was applied to growth trajectory analysis by dividing each capability and application by growth level for science teachers sharing knowledge based on subject contents.

However, the research discussed above only proposed the concepts of trajectory mining, such as finding pattern growth, identifying constraints of growth, and applying these to the trajectory and application. Therefore, to continue these concepts, the evaluation of how to apply the technique to real data is required. Learning plans are essential to students because through good plans and recommendations, students can achieve their academic goals more easily and successfully, and can also learn more effectively.

To guide students in a study plan, this study proposes a function that generates planning support for students based on their growth history. The intent of this plan is to recommend study areas and assist students in planning their studies. This paper consists of three parts. First, we will present spiral learning on capability space. Second, we will introduce the planning support functions. Third, we will provide an example of planning support for students.

2 Spiral learning on capability space

2.1 Term Definition

Let us define first the capability space and capability structure by assigned capabilities C_i: $C_1, C_2, C_3, \ldots C_n$. Each capability C_i has some levels: l_{i1}, l_{i2}, l_{i3},.., . Generally, each has five levels: l_{i1}, l_{i2}, l_{i3}, l_{i4}, l_{i5}. Unit U_i represents a growing object such as a student or teacher. For example, a student in English class seeks to improve capability such as listening capability, speaking capability, etc. We can prepare five levels: low, fair, good, very good, and excellent.

2.2 Visualization path of growth trajectory

A visualization path is a path of each capability level represented by a vector type. Each path can reach a destination or goal depending on individual's current capability level. We can create a visualization path by analyzing data from a student's growth log or that from a questionnaire if log data is unavailable or for a new student.

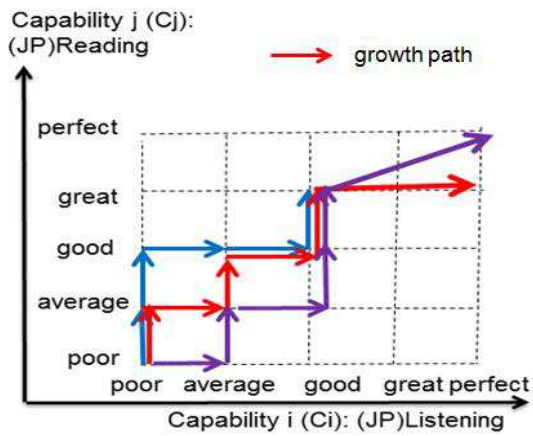


Figure 1. Example of growth trajectory

As shown in Figure 1, the growth trajectory indicates the path that can be followed to increase a capability by visualization of the path and interrelations between capabilities. In the example, C_i represents listening capability and C_j is reading capability. There are five achievable levels (poor, average, good, great, and perfect). Three possible paths are shown, indicating ways to improve their capability levels in the two capabilities.

We now define the capability space. The basic idea comes from evaluation of knowledge space theory [8], [9] for knowledge acquisition. Then, capability space is a set of potential capabilities. Furthermore, we define capability subspace as a set of state vectors that exist within two or more capability axes. We take two axes C_i and C_j, for example. Figure 2 shows an example where $X_i(t)$ is a capability state for unit U_i at $X_i(1) = (l_{11}, l_{21}, l_{31} \ldots l_{n1})$. We assume units grow only by one capability for one change from state (l_{i1}, l_{j1}) *to* (l_{i+1}, l_{j+1}).

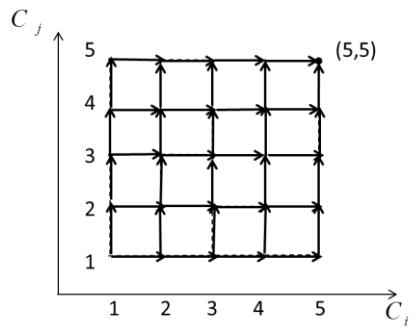


Figure 2. Two-axis based Capability Space

A capability structure suggests dependency between two . For example, high level such as l_{i5} in Figure 3(a) can be achieved without any pre-condition imposed by C_j or vice versa. However, achieving level l_{i5} as in Figure 3(b) may sometimes require satisfying another condition. In this example, l_{i5} is achieved only if a unit achieves l_{i4} first, but l_{i4} requires that l_{i3} and l_{j3} are met first, thus l_{i3}, l_{j3}, and l_{i4} are prerequisites to achieving l_{i5}.

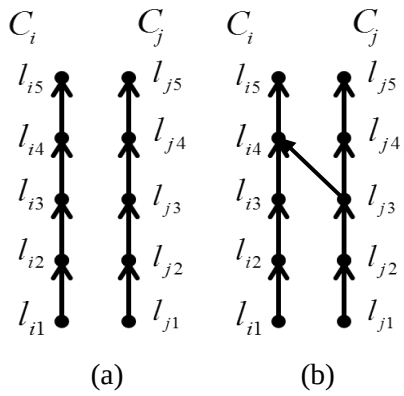


Figure 3. Example of capability structure

2.3 Plan–Do–See Cycle

Plan–Do–See is a principle applied to improve the quality of general management [10] (Figure 4). "Plan" is preliminary analysis and prediction of effects that may occur on account of upcoming

changes. For example, “Plan” is to pass the Japanese language proficiency N1 test within five years. “Do” is what is being done and there is little to change in plan control. For example, a diligent student reads and practices Japanese all the time with the intent of passing the N1 exam as planned. “See” is the evaluation work that has already been done and revised to improve the student’s chances of meeting the goal.

For example, a student has done as planned, then it is determined whether the N1 has been passed or not. If the goal has not yet been achieved, the plan is revised. Thus, planning support is based on the Plan–Do–See cycle.

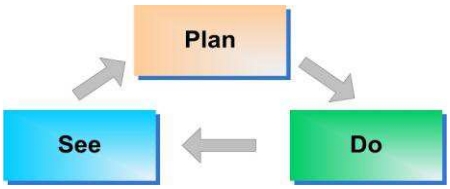


Figure 4. Plan–Do–See cycle

2.4 Review of previous systems

The previous research [5] presented the development of a prototype system for supporting growth by using visualization of the growth history. The research was divided into three main parts: first was the input of capability data for students. For this part a dataset (learning history data) was imported into the system. The second part was to visualize the learning trajectory based on the learning history data from the first part. Visualization included two types: individual and group visualization. The last part was the analysis of history data in order to recommend learning paths to a student.

This system would require extensive manual intervention in order to analyze the results for multiple students, thereby requiring a large amount of analysis time and cost. Moreover, the system only supports the “See” part of the Plan–Do–See cycle. Table I shows the comparison between available functions of the process discussed in [5] and this research. Visualization function support supported only the evaluation purpose (See). FMG (finding message generator) planning shall support the “Plan” function. The proposed planning support function supports both “Plan” and “See” functions in the PDS cycle.

Table I Comparison of functions

No	Function	Plan	Do	See
1	Visualization	-	-	●
2	FMG	●	-	-
3	Planning support	●	-	●

Planning support function will support “plan” and “See” because the process of extracting the planning support function must have to plan from student growth log. Then, we will get the extracted data as the planning result (See Figure 7). For “See”, the result of extracted planning support function will show the result from the past until the future which students have been recording. This means the follow up and evaluation from planning.

3 Planning support function

3.1 Basic principles

In this study, we use the basic principles of CMMI that are used to evaluate software organization through analysis of software capability and maturity. Both maturity and capability are indicators of consistency and growth trends reflecting the quality of a software product. CMMI applies growth levels to maturity and capability from level 1 to level 5, as shown in Figure 5.

Figure 5. Capability Maturity Model Integration

In addition to CMMI, we introduce the Spiral Enhancement Capability Support (SPICE) system [11], as shown in Figure 6. SPICE is a learning support system that can store and update student growth logs. In this paper, we use SPICE to store and analyze logs for growth trajectory and capability structure planning. SPICE also allows users to share experiences with each other. Moreover, SPICE allows a supervisor to improve student capabilities through effective planning.

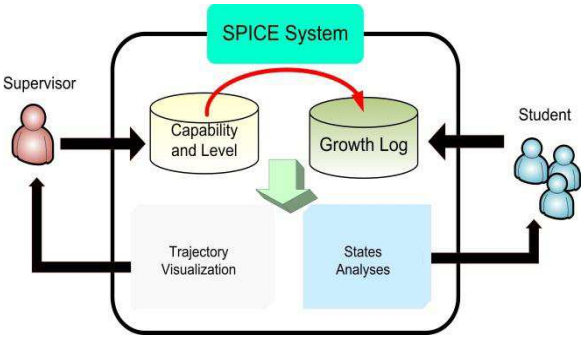


Figure 6. Model of SPICE

3.2 Process for planning support function

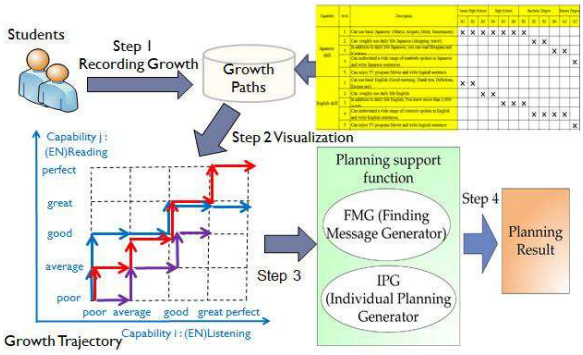


Figure 7. Finding message process

Let us divide the process of the planning support function into four steps. The first process is to record student histories. This process collects data through student questionnaires in some cases (if documented history data is not available). Students provide achievement data per year reflecting their current capability levels.

The second process is to build a visualization path for each capability level. Each path can achieve the planned destination depending on an individual's current capability level.

The third step involves generating the planning support function. This step consists of two functions. (1) FMG is the extraction message from student learning histories. The algorithm used to generate messages is Interpretive Structural Modeling (ISM) [12]. This section focuses on the finding message for dependencies between capability pairs. The details of the finding message for dependency capability pairs will be discussed in the next section. (2) Individual Planning Generator (IPG) is the plan generated from extraction of a student's learning history. The process of extraction is explained in Section 3.4.

Finally, the last step provides overall results. One result comes from the FMG function to recommend enrolling students in the types of student groups they will benefit from. The result from the IPG function will recommend an individual study plan.

3.3 Flowchart for generating a message

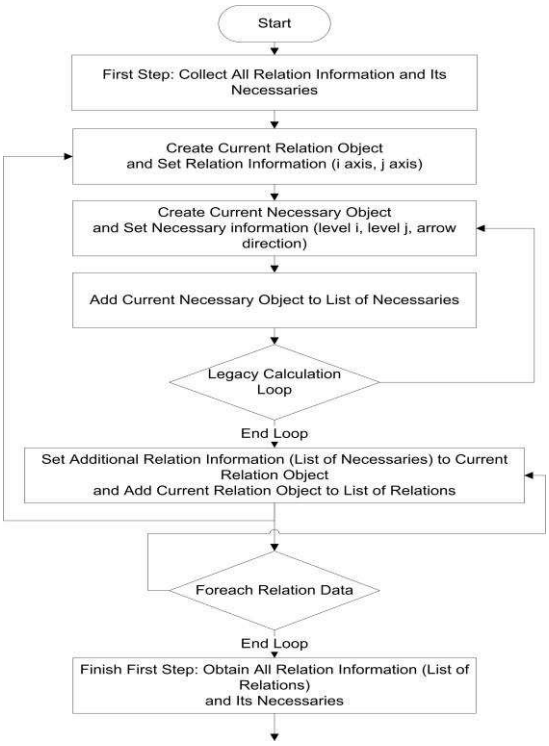


Figure 8. Flowchart for generating a message (Step f1).

FMG follows three steps as follows:

Step f1, (Figure 8) we collect the values of Relation and Necessary by setting all capability pairs in the object class. For this process, we use Loop (Legacy Calculation Loop) to calculate the values of Necessary. Then, we create Object class, which is the agent for Relation, and then collect all Relations' data (Axis i, Axis j, Necessary) in that page to create a list used in Step f2.

Step f2, (Figure 9) Searching for the Strong Dependency Relation and no Dependency Relation.

Searching for Strong Dependency: As the first step in searching for Strong Dependency values, we define the variable Max Relation Count for determining the maximum number of relationships of type Relation. An array will be packed

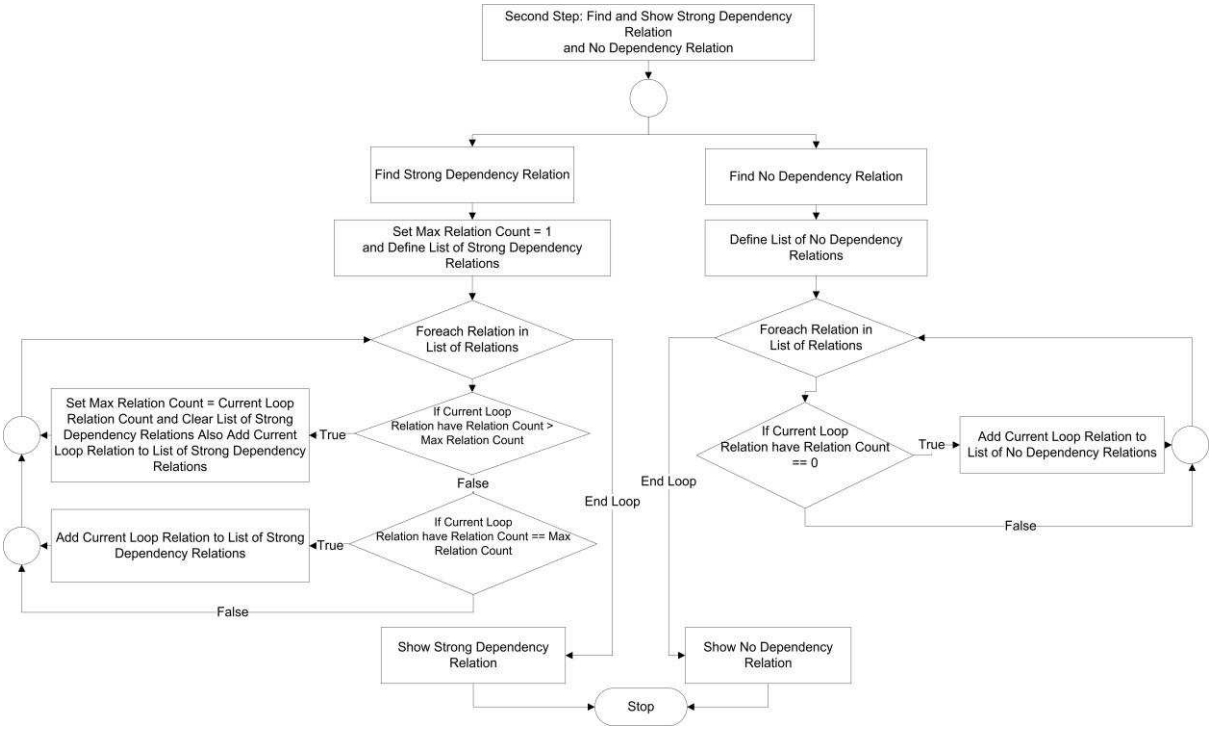


Figure 9. Flowchart for generating a message (Step f2).

only with Strong Dependency Relation values (the Strong Dependency Relation Array). We begin by setting Max Relation Count equal to 1. After that, we perform the Loop to evaluate each Relation obtained from the previous step.

We determined that the value of relationships of Relation in the Loop is higher than that of Max Relation Count, and then perform the following three steps:

1) Increase Max Relation Count to equal to the number of relationships of Relation in the Loop.
2) Clear the Strong Dependency Relation Array.
3) Insert the Relation value in the Loop into the Strong Dependency Relation Array.

If the value of Relation's relationship in the Loop is equal to Max Relation Count, it inserts the Relation value in the Loop into the Strong Dependency Relation Array. If the value of Relation's relationship in the Loop is lower than Max Relation Count, no action is taken.

Searching for No Dependency: When searching for No Dependency, we apply a principle similar to that for searching for Strong Dependency, except that, in this case, we determine whether the value of each Relationship's Relation is equal or not equal to 0. If the number is equal to 0, we place the Relation into the No Dependency Relation Array.

Step f3 simply displays the Strong Dependency and No Dependency. We apply the Strong Dependency Relation Array and No Dependency Relation Array as display data.

3.4 Flowchart for individual planning generator

We describe the flowchart for the IPG function in this section. The process begins by reading output data from the visualization function. Then, we set advice or recommendation type to create planning support. Later, the student status is checked.

The process has three conditions that must be checked; the first condition checks whether a student is over-focused on one side of the capability structure, and if so, the result will recommend that the student should get back to studying other skills first. The second condition determines if a student can or has completed two capabilities in level 5, the result is "You are excellent," etc. The third condition check is to determine if a student's data shows that both capabilities have not yet been completed at level 5, and then the process will check current student levels.

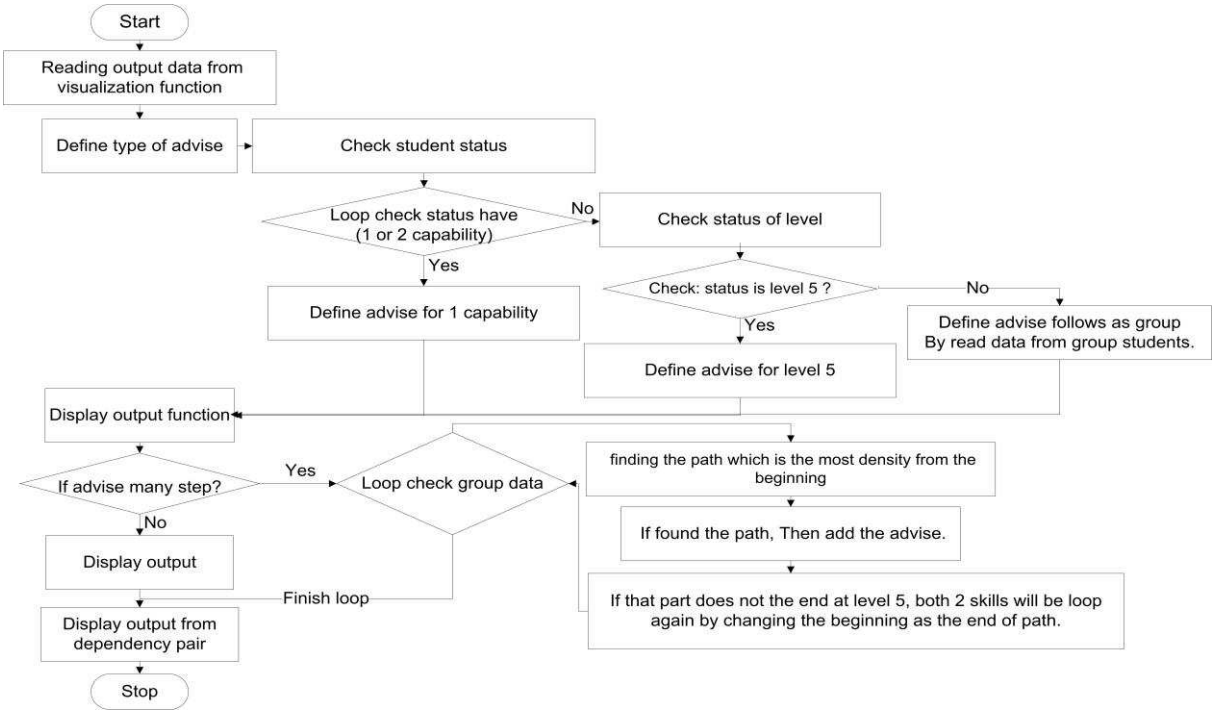


Figure 10. Flowchart of individual planning generator.

The process loops to check the conditions searching for paths that have the largest number of students and will choose that path and provide planning support advice for each path. Functions will continually check until the student's status shows completion at level 5 in both skills. Finally, it will display the output from the IPG function with the FMG function as stated above.

4 Example of planning support

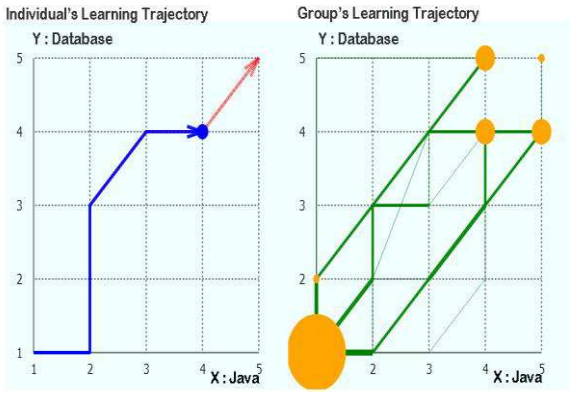
4.1 Sample Data

As a case study, we use samples of students from the Thai-Nichi Institute of Technology. The data samples consist of the following: undergraduate Information Technology (IT) students in third and fourth year, undergraduate Japanese Business students in the third year, undergraduate Computer Engineering students in second year, and graduate level IT students.

A total of 60 students are represented in the case study. We define eight capabilities (*English skill, Japanese skill, presentation skills, Java skills, study planning, web skills, e-mail skills, and database skills*). Each capability is assigned a capability level so that current student capabilities can be evaluated and study plans can be recommended.

4.2 Example plan

We randomly select one student to show an example of a study plan. We use the *Java skill* and *database (DB) skill* as the capabilities.



Advice (Planning study)
You should achieve Java skill level 5 first.
Then, you should achieve DB skill level 5

For student group of Java and DB
You should achieve Java skill level 3 before DB skill level 5.

Figure 11. Planning support example

In Figure 11, the student is currently between DB skill level 4 and Java skill level 4. In this example, the student must achieve the Java skill level 5 in order to have a chance to achieve the DB skill level 5 at first. In addition, when the characteristics of all students are considered, a minimum Java skill of level 3 is required to reach a DB skill of level 5. From the group sample, a general statement can be made that students need to study and pass the test of Java skill level 3 before they can take the skill level 5 DB test.

5 Conclusion

We have presented an automated function of planning support for students based on trajectory mining and have also presented a case study example. The goal of this study is to provide a method to guide students to select a study plan to further their academic growth. This study proposed the process of a planning support function comprising four steps, as previously discussed. We chose one student out of 60 to demonstrate a study plan to achieve success in two capabilities and showed how the plan depended on the interrelationship between the two capabilities.

The current proposal focuses on planning support functions for students involving two functions. The FMG function is a part of two planning support functions. This study illustrates the finding message of dependency only for capability pairs. The result may be not enough to create a study plan for student groups in which capabilities with multiple dependencies exist. In the future research, we will investigate other methods to accomplish the finding message function in order to guide students with effective group-tailored study plans.

Acknowledgement

The authors would like to sincerely thank the students of TNI who responded to our survey. This work was partially supported by Japanese KAKENHI-A (13370017) and KAKENHI-C (2350049).

References

[1] H. Tsuji, R. Saga, A. Tiwana and D. Albert. Trajectory Mining on Capability Space Its Concept and Potential Application. *Proc. of Forty sixth Hawaii International Conference on System Sciences*, 3705–3713, 2013.

[2] C. Romero, S. Ventura and E. Garcia. Data mining in course management systems: Moodle case study and tutorial Science Direct. *International Journal of Computer & Educations*, 51(1): 368–384, 2008.

[3] T. Kaewkiriya, R. Saga, T. Ichinotsubo, S. Aoki, W. Niu and H. Tsuji. Students Capability Growth Trajectory Mining. *Proc. of International Conference on Japan Society for Artificial Intelligence*, 2013 (Accepted).

[4] Y. Nakamura, W. Niu, T. Ichinotsubo, R. Saga and H. Tsuji. Capability Structure Extraction from Growth Trajectory. *Proc. of The paper of Technical meeting on Information System*, 33–44, 2012.

[5] T. Ichinotsubo. Develop a Prototype System for supporting Growth Using Visualization of the Growth History. *Graduation Thesis of School of Engineering at Osaka Prefecture University*, 2012.

[6] M. Paulk, C. Weber, B. Curtis and M. Chrissis. The Capability Maturity Model: Guidelines for Improving the Software Process, *Addison-Wesley*, 1995.

[7] M. Sakoda, Y. Wada, H. Tsuji and K. Seta. Social Network Service with Maturity Level for Science Teachers. *Proc. of International Conference on System, Man and Cybernetics*, 1718–1723, 2009.

[8] D. Paul and F. Jean. Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23(2): 175–196, 1985.

[9] I. Masahiro. Analysis of Information Table by Rough Set Theory. *International Journal of System control and information*, 49(5): 165–172, 2005.

[10] W. Shewhart and W. Deming. PDCA Cycle. http://en.wikipedia.org/wiki/PDCA, (Access on May 2013).

[11] Y. Nakamura and H.Tsuji. Spiral Capability Enhancement Support System. *Proc. Of International Symposium on Aware Computing*, 233–238, 2010.

[12] J. Warfield. Developing interconnected matrices in structural modeling. *IEEE Transactions on Systems Men and Cybernetics*, 4(1): 81–87, 1974.

Detecting Outlier in Graph Structure Data Using Centrality

Yukihiro Takayama† Ryosuke Saga† Takao Miyamoto†

†Graduate School of Engineering Osaka Prefecture University, Japan
{takayama@mis., saga@}cs.osakafu-u.ac.jp
aki@center.osakafu-u.ac.jp

Abstract

This study describes an outlier detection technique for graph structure data that uses the centrality index. Existing techniques set thresholds for link and node regularity. However, existing techniques are not objective and do not apply to data without the link strength information. Therefore, we pay attention to centrality, which is an index used in network analysis. We perform outlier detection objectively and on the basis of a fixed rule by using centrality. Using the proposed technique on two kinds of data, results show that unifying two or more centrality indexes for outlier detection is the most useful approach.

Keyword: data mining, outlier detection, graph structure data, centrality

1 Introduction

In recent years, with the development of information-based societies, large quantities data are being stored. Therefore, data mining is an attractive way to extract knowledge from data as well as analyze the data and perform analysis and prediction [1][2][3].

Data mining is performed according to KDD process [4]. One of the important processes in KDD is outlier detection. Outliers are data points that differ from the rest of a data set and affect the reliability of the analysis results [5][6]. Generally, we use outlier detection techniques such as DBSCAN and STING [6][7] on data that has certain attributes, such as a name or type.

However, graph structure data, which only has only one node and one link, does not contain attributes. For example, the data showing human relations on Twitter is graph structure data [8]. Therefore, outlier detection techniques such as DBSCAN and STING cannot be used for graph structure data; therefore, the outliers are often filtered on the basis of a threshold set by subjective criteria.

In this study, we propose an outlier detection technique using the centrality index [9] for graph structural data. Centrality is used to measure the importance of a node. By using centrality, we try to detect outliers in graph structure data.

This paper is organized as follows. In Section 2, we describe the related work in outlier detecting methods. In Section 3, we describe centrality. In Section 4, we introduce the algorithm used in the proposal technique. In Section 5, we show a verification experiment of the proposal technique. In Section 6, we describe our conclusions.

2 Related Work

In general, data mining aims to locate knowledge within a dataset using the KDD process. This process consists of four steps; data cleaning, data selection, data patterning, and knowledge discovery from the processed data.

Data patterning research has tended not to focus on data cleaning, which refines data in the first stage of the KDD process. However, the data cleaning step has received more attention in recent years. Although patterning is important for knowledge discovery, data patterning is not effective if the data cleaning is not performed efficiently. One data cleaning technique is outlier detection, which is the focus of this paper.

An outlier is a data point that differs from the rest of the data [5][6], as mentioned in Section 1. For example, in Japan in 2011, the National Tax Agency site reported that the average annual salary of minors was 1,060,000 yen. In this situation, an individual earning 10,000,000 yen would be regarded as an outlier. If numerous outliers such as this are included in the data set, we cannot accurately analyze the data.

Thus, it is necessary to detect the outliers that exist in data by mining the data set. Removing outliers makes it possible to analyze the data more simply and effectively.

Outliers that are visible to the naked eye (such

as missing values) can be removed by using SQL in a database. However, some outliers cannot be easily identified. It is necessary to use various techniques such as DBSCAN and STING [6][7] to analyze and remove these outliers.

The DBSCAN and STING techniques plot data to coordinate the space on the basis of attributes and detect outliers on the basis of the distribution. DBSCAN detects outliers on the basis of distance between the center of a cluster and the individual data. STING decomposes the spaces into grids and detects the data outliers in the grid by locating low density areas. That is, STING techniques consider the data separated from the center of the distribution to be outliers. However, because these techniques utilize a data distribution, they are not applicable for graph structure data.

Existing research attempting to detect outliers in graph structure data set a fixed threshold in the node and the link on the basis of a researcher's experience or trial and error. For example, Saga et al. filtered links and nodes experimentally by using the Jaccard Coefficient [10][11]. By preparing a threshold, isolated nodes and links were considered outliers and deleted. The idea is that important nodes and links are strongly related to other nodes and links, and thus have numerous connections with others and are rarely isolated.

However, because the threshold is set subjectivity, a correct desired result cannot be obtained by confusing process. Moreover, a technique that sets a link threshold cannot apply to the data without the information regarding the strength of a link. As a solution to these two problems, we also pay attention to the link's centrality index, which is a measure of network analysis.

In this study, we propose a method that detects outliers on the basis of certain objective rules using centrality for graph structure data. This proposal technique is performed on graph structure data from an undirected graph.

3 Centrality

Centrality is one of the indexes used in network analysis [8]. It is used to compare the importance of each node in a network and the relation between nodes. For example, in a network containing certain human relations, calculating a certain person's centrality shows the position in the human relation or the degree to which that person is trusted by the surrounding people. Moreover, centrality can show the popularity of a Web page or what kind of page it is.

Centrality can be determined regardless of a node's attribute. Centrality is not directly reflected in the attribute of a node. For example, in a network containing certain human relations, attributes such as a person's character and evaluations from friends do not reflect centrality. Although those may increase the number of links to a node, centrality is determined by the network structure.

Various kinds of centrality have been proposed. The following is an explanation of the features and calculation methods in the three kinds of the centrality that we use in this research.

3.1 Degree Centrality

A degree is the number of links that are connected to a node. The degree centrality shows that the more links there are connected to a node, the more likely it is a center node. If the adjacency matrix of a certain graph is set to a_{ij}, the degree centrality of node i is formulated $C_d(i)$ as follows:

$$C_d(i) = \sum_{j=1}^{n} a_{ij} \tag{1}$$

3.2 Closeness Centrality

The closeness centrality shows that the closer a node is to other nodes, the more likely it is a center node. The distance is the shortest path from a certain node to other nodes. If the sum of the shortest path from node i to other nodes is set to $\sum_{j=1}^{n} d_{ij}$, the closeness centrality of node i is formulated $C_c(i)$ as follows:

$$C_c(i) = \frac{n\text{-}1}{\sum_{j=1}^{n} d_{ij}} \tag{2}$$

3.3 Betweeness Centrality

The betweeness centrality shows that the more a node is in a path that connects different nodes, the more likely it is a center node. For example, it is important that a node is in the path that connects two networks. If the number of the shortest path from node j to node k is set to $g_{jk}(i)$, and the sum of the paths that pass node i in the shortest paths from node j to node k is set to $C_b(i)$, the

betweeness centrality of node i is formulated B as follows:

$$C_b(i) = \sum_{i \neq j \neq j}^{n} \frac{g_{jk}(i)}{g_{jk}} \quad (3)$$

4 Outlier Detection Using Centrality

4.1 Basic Idea

Existing techniques have two issues: setting a subjective threshold may not result in correct desired results and it is appropriate to set a threshold on the basis of a link's strength if the data does not contain link strength information. We use centrality to address these two issues.

An important node is located in an important place in a graph structure and the centrality of such a node is high. Such a node is also closely located to other important nodes. Conversely, a node that has relation with other sparsely positioned nodes is an unimportant node. Such a node is located in an unimportant place in the graph structure and thus centrality of such a node is low.

We considered that the number of nodes is plotted to table to elucidate rise and fall of value. Then we considered that outliers are detected objectively by paying attention to the point of centrality with the largest reduction in the number of nodes. This is why we consider that a large reduction in the number of nodes indicates a set of nodes with a small centrality, and that such a set included many outliers because outlier is small centrality. So I considered being able to detect outliers by cutting nodes that are below centrality of large reduction.

With the aforementioned assumption, we propose the following outlier detection method, which consists of Steps 1 through 3.

Step 1 Calculating Centrality

In Step 1, we calculate three kinds of centrality: degree, closeness, and betweeness. The calculated centrality is normalized between 0 and 1.

Step 2 Grasping Outliers

To make it easy to compare the distribution of the number of nodes in each centrality, we create a histogram, where the vertical axis is the number of nodes and the horizontal axis is the centrality. Figure 2 shows an example of the frequently table.

When we create a histogram, the centrality on the horizontal axis is divided into 10 units because it is difficult to divide up the outliers if the division interval is too large or small. We determined that 10 units was a suitable interval on the basis of our experiments.

When a range is used that includes a centrality bias, many nodes that are not outliers are contained in the range in addition to the outliers. In such an instance, it is necessary to analyze the data in detail. Therefore, if the data contains a range with 80% or more outliers, the range is then divided by 10 and is analyzed.

In this histogram, we detect node sets with a low centrality of outliers.

Step 3 Detecting Outlier

We create a rule that focuses on centrality with the largest reduction of the number of nodes and detect nodes as outliers.

4.2 Detecting Outlier Using Ensemble Centrality

Ensemble learning [6][12] is the method for creating a high precision learner by combining weak learners. We apply this method here.

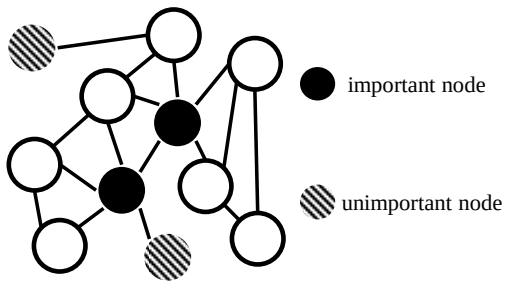


Figure 1. Importance of Nodes in Graph

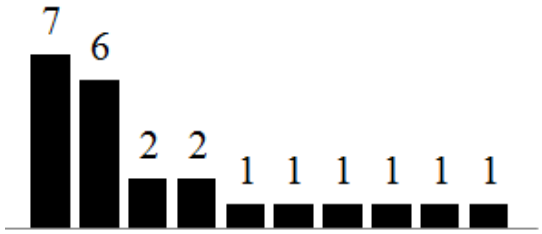


Figure 2. Example Histogram

A certain centrality may be unable to detect outliers well because we calculate each centrality one by one. Therefore, to compensate for each bad part and increase the accuracy of detection, we regard nodes that are detected as outliers in all types of centrality as outliers.

5 Verification Experiment

5.1 Environment of Verification Experiment

To verify the usability of our proposed methods, we performed experiments. We assumed that the real data was correct and generated experimental data by adding random nodes and random links as outliers. As real data, we used the "Les Miserable" and "Polbooks" data sets provided from the site [13]. Les Miserable is a weighted graph whereas Polbooks is not. Table 1 shows the details of data. We investigated the graph structure of the Polbooks and Les Miserable data and determined that the graph structure of these data was of scale-free.

We used four evaluation indices: recall, precision, F-measure, and recovery. In general, these indicators are used in examining the accuracy of detection results. Recall indicates the completeness of the detection result. Precision indicates the suitability of the detection result. F-measure indicates harmonic mean recall and precision. F-measure is high, it means that the detection result well. Moreover, we used an evaluation index: recovery. This index indicates whether the method restored original data much.

The calculations formulas of the evaluation indices are described below.

$$\text{recall} = \frac{\text{detected outlier}}{\text{added outlier}} \tag{4}$$

$$\text{precision} = \frac{\text{detected outlier}}{\text{detected node}} \tag{5}$$

$$\text{F-measure} = \frac{2*\text{recall}*\text{precision}}{\text{recall} + \text{precision}} \tag{6}$$

$$\text{recovery} = \frac{\text{undetected outlier of real data}}{\text{node of real data}} \tag{7}$$

5.2 Process of Verification Experiment

We used three processes in the verification experiments.

Table 1. Experimental Data Details

	Polbooks	Les Miserables
Nodes	104	77
Links	441	254
Diameter	6	4
Average Degree	8.42	6.60

I. We added random nodes and links as outliers to the real data.
II. We performed the proposed technique on the experiment data and evaluated whether the detected nodes were outliers.
III. We repeated steps I and II five times and calculated an average for each evaluation index.

There are five patterns in adding outliers: one time, three times, five times, seven times and nine times. We added nodes that were rounded off by 10%.

For the strength of the random outlier links, we configured them as more than one and less than the random value, which was rounded off as the average degree of data. Because the strength of a random link connecting to another outlier seems low, the strength of the random links is calculated from the uniform distribution in the first quartile of the strength of a link in the real data. If the graph was not a weighted graph, we configured the strength of the all links as 1.

5.3 Results of the Verification Experiment

Figure 3 shows the results of the verification experiments. The horizontal axis is the increase in the number of random nodes and links.

Figures 3(a) and 3(b) show that when using degree centrality, recall was always 1. Moreover, the precision and F-measure were between about 0.1 and approximately 0.7 as the outliers increased. The recovery was between about 0.2 and about 0.35.

Figures 3(c) and 3(d) show that when using closeness centrality, recall was always approximately 1. Moreover, precision and F-measure were much the same as those of degree centrality.

Figures 3(e) and 3(f) show that when using centrality, recall was between about 0.3 and approximately 0.8 as the number of outliers increas-

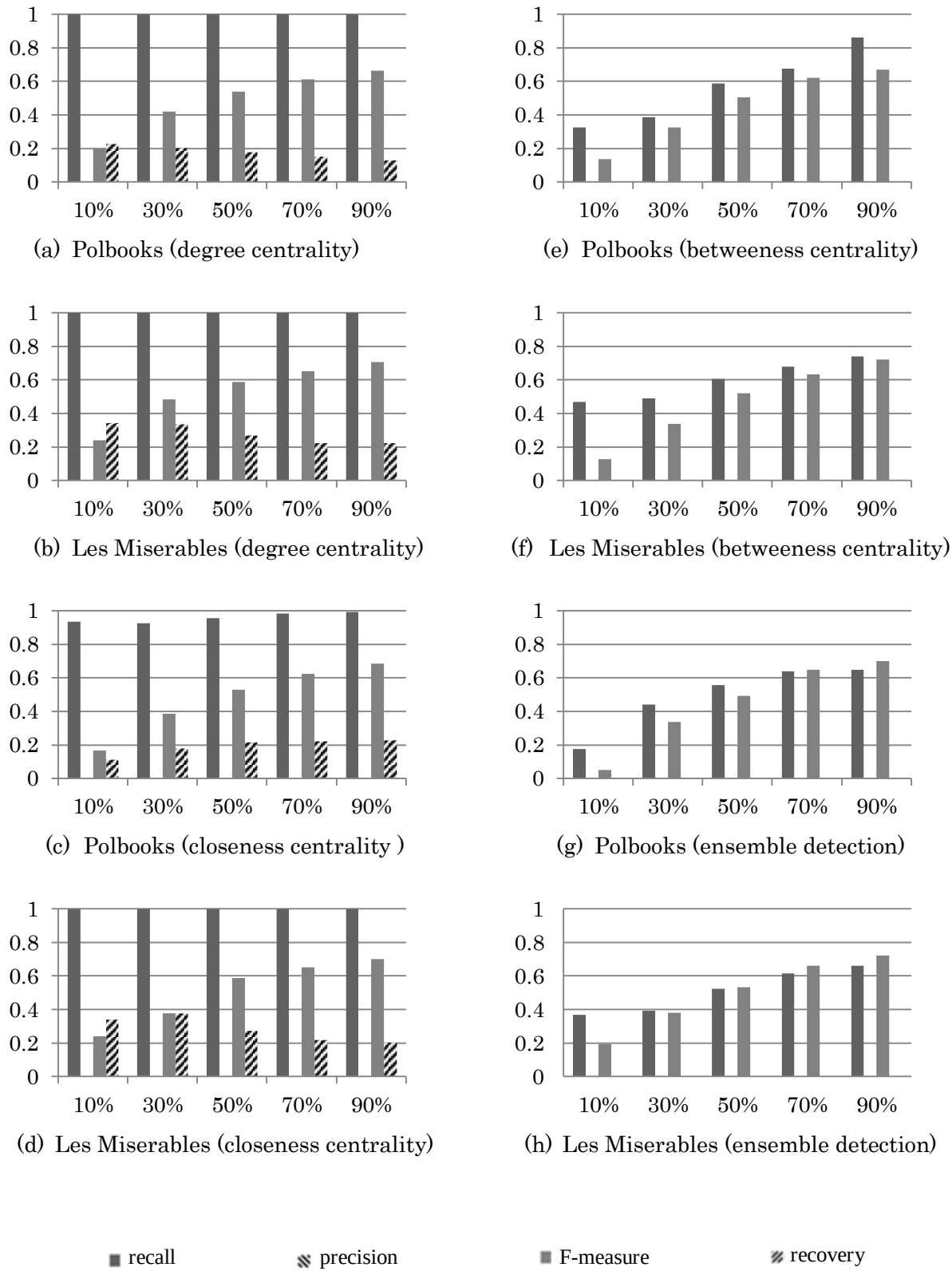


Figure 3. Results of the Verification Experiment

ed, unlike the result of the degree centrality and closeness centrality experiments. The precision and F-measure were between approximately 0.05 and approximately 0.8 as the outliers increased. The recovery was between approximately 0.4 and approximately 0.85 as the outliers increased, unlike the results of the degree centrality and closeness centrality experiments.

Figure 3(g) and 3(h) show that when using ensemble detection recall, precision, and F-measure became the same as those of betweeness centrality. However, the recovery measures between approximately 0.7 and approximately 0.85 differed from those of the degr ee, closeness, and betweeness centrality. The recovery result was best in the proposed techniques, and the ensemble detection technique best protected the real data.

5.4 Discussion of Verification Experiment

In the experiment using degree and closeness centrality, the recall results showed all the added outlier were detected. The result of precision and F-measure showed the accuracy of detection improved because the rate of the outliers in the detected nodes increased with the increasing number of random nodes. The recovery results were poor because the false-positive rate (mistaking real data for outliers) was high.

However, the results showed that the accuracy of precision, F-measure, and recovery was poor when the number of added outliers was low. We believe this is because there were more correct nodes than outliers in the range of the largest reduction of the number of nodes when the data had only a few outliers.

In the detection using betweeness centrality, the recall showed that not all added outliers were detected. The precision and F-measure exhibited the same tendencies as those of degree and closeness centrality. The recovery showed that the outliers can be detected without damaging real data with an increase in added outliers.

In the experiment using ensemble detection, the recall showed that not all the added outliers were detectable. The precision and F-measure exhibited the same tendencies as the results of betweeness centrality. This is why the result of ensemble detection is affected by betweeness centrality. The F-measure for both betweeness centrality and ensemble detection is by the more touch. However, recovery using betweeness centrality and ensemble detection improved. Therefore, we consider the ensemble detection the most useful.

In summary, in the experiments using degree and closeness centrality, although the F-measure was good, the accuracy of the recovery is poor. Therefore, it is determined that the proposed technique using degree centrality is not useful. In the experiment using betweeness centrality, the F-measure and recovery became good, and there were few instances where the real data were identified as outliers. Therefore, the proposed technique using between centrality is useful.

In the experiment using ensemble detection, the result of recovery was the best of the proposed techniques. Therefore, detection using ensemble detection is the most useful. However, the detection accuracy was poor when there were few added outliers. This problem will be addressed in future work.

From the result of this verification experiment, the same result was able to be obtained from two kinds of data in each detection. Therefore, this proposed technique can be applied to data without link strength information.

6 Conclusion

In graph structure data research, researchers detect outliers by setting thresholds for the nodes or links using only a researcher's experience. However, this approach has two issues. First, a correct desired result cannot be obtained because the threshold is subjective. Second, a link threshold cannot apply to data without the information regarding the strength of a link.

Therefore, we incorporate a centrality index to detect outliers in data that do not contain information on the strength of the links. Moreover, we detect outliers objectively by using centrality and fixing a rule in outlier detection. In the proposed technique, we detect outliers using three centrality measures— degree centrality, closeness centrality, and betweeness centrality—and conduct an ensemble detection by combining the results of the three centralities.

We validated the usefulness of the proposed techniques to data with and without link strength and using the defined evaluation index. The results of the verification experiments showed that the most useful outlier detection method used

ensemble detection.

For the two problems of outlier detection in graph structure data, one could be solved by setting a threshold on the basis of a fixed rule, and the other could be solved by using centeredness in the proposed technique. Moreover, in the results of the verification experiments, precision was poor in the case of only a few outliers. Therefore, future work will involve raising the accuracy of outlier detection in data with only a few outliers.

References

[1] B. K. Bhardwaj and S. Pal. Data Mining: A Prediction for Performance Improvement Using Classification. *International Journal of Computer Science and Information Security*, 9(4):136-140, 2011.

[2] J. Srivastava, R. Cooley, M. Deshpande and P. Tan. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations*, 1(2):13-23, 2000.

[3] R. Ng and J. Han. Efficient and Effecyive Clustering Methods for Spatial Data Mining. *Very Large Databases*, 144-155, 1994.

[4] U. Fayyad, G. Shapiro and P. Smyth, The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communications of the ACM*, 39(11):27-34, 1996.

[5] C. C. Aggarwal and P. S. Yu. Outlier Detection for High Dimensional Data. *ACM SGMDO Record*, 30(2):37-46, 2001.

[6] J. Han, M. Kamber and J. Pei. *DATAMINING Concept and Techniques Third Edition*. Morgan Kaufmann, United State, 2011.

[7] M. Ester, H. Kriegel, J. Sander and X. Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Knowledge Discovery and Data Mining*,226-231, 1996.

[8] M.J.Welch, U. Schonfeld, D. He and J. Cho. *WSDM' 11*, 327-336, 2011.

[9] M. E. J. Newman. *Networks An Introduction*. Oxford University Press, United Kingdom, 2010.

[10] B. Miraz, B. Keller and N. Ramakrishnan, Studying Recommendantion Algorithms by Graph Analysis. *Journal of Intelligent Information Systems*, 20(2):131-160, 2003.

[11] R. Saga, M. Letelier, N. Kaisaku, Y. Takayama and H. Tuji. Knowledge Discovery in Web Access Log of E-commerce Site with FACT-Graph and Sequential Probability Ratio Test. *Frontiers in Artificial Intelligence and Applications*, 243:2080-2089, 2012.

[12] M. Girolami. *Advances in Independent Component Analysis*. Springer, Germany, 2000.

[13] M. Newman. Network data. http//www-personal.umich.edu/~mejn/netdata/, (Access on May 2013).

A Robust Privacy-Preserving Recommendation Algorithm

Alper Bilge **Ihsan Gunes** **Huseyin Polat**
Computer Engineering Department
Anadolu University, Eskisehir, Turkey
{abilge,ihsang,polath}@anadolu.edu.tr

Abstract

Privacy-preserving collaborative filtering schemes are key recommender system technologies for e-commerce field. They focus on alleviating information overload problem by providing personalized recommendations without deeply jeopardizing customers' privacy. Like their non-private versions, privacy-preserving recommendation methods might be easily subjected to profile injection attacks for manipulating produced recommendations in favor or disfavor of certain products. Clustering-based prediction schemes have shown to be effective in distinguishing bogus profiles from genuine ones; and they are relatively more resilient than memory-based methods against shilling attacks in non-private environments. Motivating from this fact, we propose bisecting k-means clustering-based privacy-preserving recommendation algorithm as a robust recommendation algorithm, which was formerly proposed as a scalable and accurate private recommendation scheme, against previously defined several private shilling attack strategies. We then investigate the algorithm with respect to robustness. Thus, we perform some real data-based experiments on a benchmark data set using six profile injection attacks. We empirically show that the algorithm performs in a robust manner with insignificant alterations in predicted values when fake profiles are injected. The reason for this phenomenon is that the algorithm is able to cluster fake profiles together.

Keywords: Shilling attacks, privacy, robustness analysis, collaborative filtering, clustering, recommendation.

1 Introduction

With increasing amount of information available in everyday life through widespread usage of the Internet services, recommender system technologies have become one of the most practical tools to filter out useful information. Collaborative filtering (CF) is the key recommendation method used in e-commerce facilities to extract products of possible interest such as movies, music CDs, books, and so on. These systems rely on preference information provided explicitly by customers or implicit inferences from web logs. They recommend products by discovering liking patterns of users relying on similarities of past rating trends.

CF algorithms can be mainly classified into three classes: (*i*) memory-based, (*ii*) model-based, and (*iii*) hybrid approaches. Memory-based algorithms are successful in terms of accuracy of predictions and representing word-of-mouth habits [1]; however, they commonly suffer from scalability issues [2]. Model-based solutions, on the other hand, derive a model from the input database off-line and utilize it to produce predictions online [3]. Although their online performance is better, it is computationally expensive and hard to update the model. Hybrid systems like employing content-based filtering along with CF process [4] or employing clustering to perform neighborhood formation prior to prediction estimation [1; 5] combine both approaches' advantages for an optimal solution.

Since recommender systems rely on obtained preference information of users, they are open to manipulations of malicious users or competing companies. Such interventions are called shilling or profile injection attacks [6], which can be very effective against traditional CF schemes [7; 8]. Shilling attacks can be categorized as *push* or *nuke* attacks according to their intent. In those attacks, the attacker creates bogus profiles to favor or disfavor a particular item and inserts them into the database as a genuine user. However, clustering-based approaches are shown to be successful in distinguishing shilling profiles from genuine ones [6; 9] because fake profiles show resemblance among themselves and clustered together.

Submission of preference information causes several privacy risks for individuals such as price discrimination, government surveillance, and so on [10]. Therefore, it has been receiving increasing attention to produce dependable predictions without deeply jeopardizing individual privacy and privacy-preserving collaborative filtering (PPCF) concept has been proposed to estimate sufficiently accurate recommendations without revealing private preferences of users. Cryptographic [11], randomization-based [12], and obfuscation-based [13] methods aid privacy protection. In our previous studies, we demonstrate that memory-based PPCF schemes are also vulnerable to shilling attacks [14; 15]. However, clustering-based PPCF schemes might be robust against profile injection attacks. Thus, we propose bisecting k-means clustering-based PPCF recommendation algorithm as a robust algorithm. We investigate it with respect to robustness under several attacking strategies.

Remainder of the paper is structured, as follows: Section 2 discusses relevant literature and explains originality of the work. Then, we explain how privacy-preserving bisecting k-means clustering-based PPCF performs and shilling attacks can be implemented to modify its outputs in Section 3. Section 4 experimentally evaluates the robustness of the algorithm against manipulations under four push and two nuke attack strategies. We finally conclude the paper by a brief discussion and present some future research directions in Section 5.

2 Related Work

CF systems date back to early nineties [16]. They produce personal predictions relying on past preferences of similar users [17]. CF is integrated as a recommender system by online vendors operating on preference data.

Individuals today consider that it might be risky to submit preference data due to privacy concerns and they refrain from submitting genuine ratings [18]. However, dependable referrals can only be produced through authentic data. Hence, researchers propose PPCF schemes to provide referrals with privacy. Initial proposals rely on distributed solutions by Canny [11], where an aggregate data is formed by users of a community using singular value decomposition. Privacy is achieved using cryptographic methods and the proposed scheme can be expanded to a fully peer-to-peer system. Since most CF systems run on central servers, today's PPCF systems are designed to perform on masked data using randomization and obfuscation methods [13; 19]. In such methods, users control their privacy by masking, perturbing, and substituting their votes. Such disguising is realized by random perturbation techniques, which allows performing aggregate algebraic operations on the collection while not disclosing individual data items.

With increasing popularity of CF systems, several attacking mechanisms arise to manipulate their outputs in favor of particular products. Dellarocas [20] inspires manipulation attacks to recommender systems, where some mechanisms are defined to avoid fraud in online reputation reporting systems. O'Mahony et al. [21] then discuss vulnerabilities of automated prediction estimation process against manipulations. They also describe the amount of information needed about the database to realize effective shilling attacks. Lam and Riedl [22; 23] analyze cost of attacks and propose that there is a relation between privacy and the value of information. Several attacking methods are proposed in the literature like random, average, bandwagon, and segment attacks [24]. Shilling attacks can be designed on either a low or high level of knowledge such as average ratings of products or general tendencies of particular users. Moreover, their effectiveness are investigated against memory- and model-based CF schemes [25]. Recently, Gunes et al. [26] survey recent research on shilling attacks and provide analysis on attack descriptions, detection methodologies, robust algorithms, and evaluation metrics.

Effects of shilling attacks on memory-based PPCF schemes are deeply analyzed and shown that they are vulnerable to manipulations [15]. Novel PPCF schemes should be proposed as robust prediction algorithms. Although bisecting k-means clustering-based PPCF algorithm is proposed to protect individual privacy while providing predictions with high accuracy efficiently [27], it might be an appropriate proposal as a robust algorithm. Thus, we investigate its robustness against profile injection attacks. Unlike the explained previous studies, we focus on robustness analysis of this algorithm against profile injection attacks and show how clustering-based approaches are resistant to shilling attacks in privacy-preserving environments.

3 Attacking Bisecting *k*-means Clustering-based PPCF Scheme

Due to the reason that recommender systems are open for public usage and therefore vulnerable to manipulations, both non-private and privacy-preserving recommendation algorithms need to have robust mechanisms to estimate predictions. However, state-of-the-art memory-based CF and PPCF schemes are not resistant to such attacks and exposed to significant shifts in predicted values. In this section, we describe bisecting *k*-means clustering-based PPCF algorithm, six attack designs for masked data, and explain how robust the algorithm is.

3.1 Bisecting *k*-means Recommendation Algorithm with Privacy

Privacy risks in online shopping services direct people not willing to provide any preference information. In order to relieve privacy concerns of users, Polat and Du [12] propose the following data masking procedure so that individual preferences are not disclosed to the server; yet, aggregate algebraic operations can still be performed and recommendation can be generated.

- The central server lets each user know global data perturbation parameters, i.e., σ_{max} and β_{max}.
- User u calculates her ratings' z-score values, chooses a distribution (uniform or normal), and creates random disguising and filling values by chosen σ_u, where $\sigma_u \in (0, \sigma_{max}]$.
- Each user then perturbs each z-score value by adding a random value onto it and fills $\beta_u\%$ of her empty cells, where $\beta_u \in (0, \beta_{max}]$.
- Finally, each user submits her perturbed ratings vector to the server.

The central server collects such disguised user vectors and forms a user-item matrix, $U'_{n\times m}$. At the beginning, the server forms a binary decision tree off-line by utilizing bisecting *k*-means clustering algorithm on U'. Given U', and an optimal value of number of neighbors (*N*), *k*-means clustering is applied to divide U' into two clusters at each level (hence, it is called bisecting) and cluster centers are indexed to be used as a forwarding tool for each corresponding level. If number of users in any cluster exceeds *N*, then such clusters are continued to be divided recursively into subsets via the *k*-means clustering. Finally, a binary decision tree is obtained having indexed cluster centers as branch nodes and grouped neighbor users at leaf nodes.

When an active user a arrives and queries for a prediction, instead of calculating similarities with all users, the server only forwards the active user according to her similarity to two cluster centers at each level. While traversing, two similarity calculations are performed at each level to determine next hope (either right or left). Finally, the leaf node that the new user belongs is determined and all users in that corresponding node and its sibling are regarded as neighbors. Then, a prediction is calculated as a weighted average of neighbors on target item is calculated as formulated in Eq. 1 and returned to the active user as a prediction.

$$p_{aq} = \overline{v_a} + \sigma_a \times \frac{\sum_{u\in N} w_{au} \times z_{uq}}{\sum_{u\in N} w_{au}}. \quad (1)$$

In Eq. 1, p_{aq} is the prediction for a on target item q, $\overline{v_a}$ and σ_a are mean value and standard deviation of a's ratings, N is the set of neighbors, z_{uq} is the z-score of user u on q, and w_{au} is the similarity between a and neighbor u. Afterwards, the tree continues growing in such a way so that if any leaf node population exceeds the stopping criterion, the server immediately bisects that leaf node to grow. Therefore, it is a continual process to update the decision tree.

An example binary decision tree produced by the algorithm is presented in Fig. 1, where initially there are 150 users in the database and the stopping criterion is determined as 20 users. Therefore, the algorithm bisects each internal node starting from the root until there remains at most 20 users in a leaf node. At each internal node, two cluster centers (C^L and C^R) are recorded to be used in forwarding purposes.

In order to attack such recommendation scheme, the attacker produces attack profiles and inserts them into the system. Clustering-based CF schemes are successful in detecting fake profiles or bogus profiles [6; 9]. Bhaumik et al. [9] shows that shilling profiles show resemblance to each other; therefore, when they are clustered, they tend to move together into the clusters. Using this intuition, clustering can be used

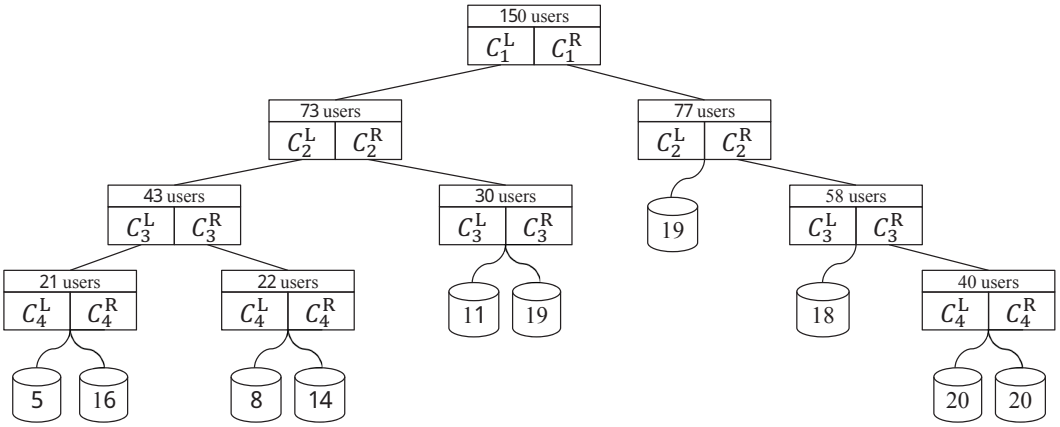


Figure 1. An example binary decision tree

as a detection tool for shilling attacks and it can be utilized to create robust recommendation algorithms. In this study, we hypothesize that such intuition works best for clustering into two clusters to move shilling profiles together. In addition, applying such clustering repeatedly is supposed to eliminate all shilling profiles after some level of the produced binary decision tree. Therefore, we claim that malicious profiles substantially distinguishes from genuine ones after some level of the tree and it becomes very unlikely for any active user belonging to a leaf node consisting of fake profiles. As a result, the proposed recommendation scheme performs robust against shilling attacks.

3.2 Shilling attacks on masked databases

Profile injection attacks are generated by inserting fake or shilling profiles into user-item databases. There are several attacking strategies in the literature, which can be explained in a general format as depicted in Fig. 2 [24]. Note that I_S, I_F, and I_ϕ refers to selected, filler, and empty cells in the fake profile, respectively; and a unique item, i_t, is targeted. Selected items are chosen for characterizing an attack, filler items are chosen to prevent easy detection of fake profiles, and the target item is assigned either a high or a low rating value for push and nuke attacks, respectively.

As part of their generic form, four push attacks and two nuke attacks covered in this paper can be described, as in the following [15]:

Random attack (RN). Selected items set is empty. Arbitrarily chosen filler items are filled with random values and the target item is assigned a high random value.

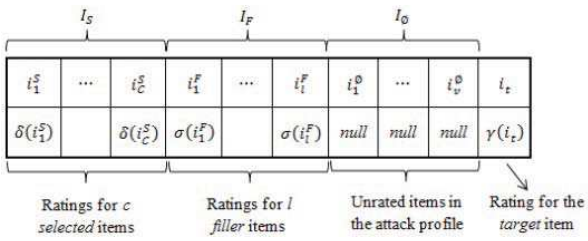


Figure 2. General form of an attack profile

Average attack (AV). Selected items set is empty. Arbitrarily chosen filler items are filled with close random values to the item's mean and the target item is assigned a high random value.

Bandwagon attack (BW). Selected items are chosen among popular items (having high means). Selected and filler items are filled with random values, where selected items are assigned the highest ones and the target item is assigned a high random value.

Segment attack (SG). Selected items are chosen from high average items with a certain property (such as horror movies or jazz music). Selected and filler items are filled with random values, where selected items are assigned the highest ones and the target item is assigned a high random value.

Reverse bandwagon attack (RBW). Selected items are chosen among unpopular items (having low means). Selected and filler

items are filled with random values, where selected items are assigned the lowest ones and the target item is assigned a low random value.

Love/hate attack (L/H). Selected items set is empty. Arbitrarily chosen filler items are filled with very high random values and the target item is assigned a low random value.

4 Experimental Evaluation

After explaining how shilling attacks can be implemented over privacy-preserving bisecting k-means recommendation algorithm, we conducted real data-based experiments to scrutinize effects of such attacks on system robustness with respect to two controlling parameters. The controlling parameters, filler size and attack size, are defined in the literature for designing effective shilling attacks. Filler size parameter indicates the percentage of cells to be filled out with fake ratings while creating the attack profiles. Attack size can be described as the pre-attack profile count proportional to the number of users in the database. We conducted two sets of experiments with varying values of the explained parameters against the bisecting k-means clustering-based PPCF scheme.

4.1 Data set and evaluation criteria

In the experiments, publicly available MovieLens (ML) data set, which is collected by GroupLens (http://movielens.umn.edu/) is utilized. It is the most widely used and well-known real collection for CF purposes. It holds 100K ratings from 943 users on 1,682 movies and the rating range allows 5-star discrete numeric values.

We used prediction shift metric in order to measure the prediction alterations due to effects of shilling attacks. Prediction shift can be described as the average change in the prediction for the attacked item before and after the attack profiles are included.

4.2 Experimental Results and Discussion

During the experiments, we followed all-but-one experimentation methodology, which enables full utilization of the data set. This methodology considers one of the users as the active user a and the rest of the set as the training users at each of the iterations. The utilized attacks target two separate sets of 50 movies for push and nuke attacks. Those sets for push and nuke attacks are constructed selecting arbitrarily from different rating ranges to represent characteristics of the original data set. Since it is unreasonable to push a popular item with high ratings or similarly nuke an unpopular item, we principally selected items with low mean values to push and high means to nuke. Table 1 shows the statistics of 50 target movies for push and nuke attacks, where each value indicates how many of the movies fall into corresponding group.

Table 1. Statistics of targeted movies

Ratings	Pushed Items		Nuked Items	
	1–2	2–3	3–4	4–5
1–50	30	15	12	18
51–150	—	3	5	6
151–250	—	1	2	3
250 and up	—	1	1	3

In the experiments, all target items were attacked individually for all users in the system. Binary decision trees were constructed by omitting and including fake shilling profiles. Then, predictions were estimated based on the produced binary decision trees and prediction shift values were observed to demonstrate relative change on predicted values for differing attack models. The stopping criterion for building binary decision trees is set to 30. Although varying stopping criterion might alter obtained prediction shift values especially with varying attack sizes, we fix such parameter due to page limitations and discuss algorithm's robustness performance relying on a constant stopping condition value. Additionally, all experiments were repeated 100 times due to randomization in perturbation process and average results were calculated. We exclusively presented the obtained empirical results for push and nuke attacks in the following sections.

4.2.1 Effects of push attack models

In order to investigate the effects of the four push attack models on bisecting k-means clustering-based prediction algorithm with privacy, we performed experiments by varying filler size in the attack profiles and varying attack size inserted into the database. Filler size parameter indicates the number of fake votes for the filler items added to fill out the attack profile; and thus, it is directly related to the success of the attack.

Attack size determines the number of inserted attack profiles; hence, it is also vital in realizing significant manipulations. To clarify the effects of these parameters, we kept each other at 15% while experimenting on one of them. In the experiments, we varied filler size from 3% to 25% and attack size from 1% to 15%. Overall results are shown in Table 2 and Table 3 with varying filler and attack sizes, respectively. The highest prediction shift values are given in bold.

Table 2. Prediction shift with varying *filler size* for push attack models

	Filler size				
Attacks	3%	5%	10%	15%	25%
RN	.011	.013	.018	.031	**.053**
AV	**.006**	.004	.003	.002	.002
BW	.019	.023	.032	.035	**.036**
SG	.195	.200	**.209**	.201	.195

Table 3. Prediction shift with varying *attack size* for push attack models

	Attack size				
Attacks	1%	3%	6%	10%	15%
RN	.015	.039	.069	**.078**	.052
AV	.004	.010	**.012**	.002	.001
BW	.020	.037	.036	.036	**.039**
SG	.037	.166	.168	.180	**.196**

As can be seen from Table 2 and Table 3, none of the attack models is able to achieve a significant prediction shift for varying filler and attack sizes. With increasing filler size and attack size, the effects get larger; however, increasing these parameters more is not feasible for the sake of detection of attacks. Only segment attack achieves a maximum shift value around 0.2, which can still be ignored because it is only one in twentieth of the prediction range. Also note that privacy-preserving bisecting k-means clustering-based prediction algorithm runs with about 0.72 mean absolute error, as reported in [27], which is about four times larger than the maximum achieved prediction shift. Therefore, we can conclude that bisecting k-means clustering-based prediction algorithm is robust against profile injection push attacks in privacy-preserving environments.

4.2.2 Effects of nuke attack models

In addition to push attack models, we also investigated two well-known nuke attack models on privacy-preserving bisecting k-means clustering-based prediction algorithm. Similarly, we performed experiments by varying filler size and attack size; and kept each other at 15% while experimenting on one of them. Overall results are displayed in Table 4 and Table 5 with varying filler size and attack size, respectively. The highest prediction shift values are given in bold.

Table 4. Prediction shift with varying *filler size* for nuke attack models

	Filler size				
Attacks	3%	5%	10%	15%	25%
RBW	-.047	-.050	-.049	**-.054**	-.037
L/H	**-.001**	.002	.003	.003	.007

Table 5. Prediction shift with varying *attack size* for nuke attack models

	Attack size				
Attacks	1%	3%	6%	10%	15%
RBW	-.018	-.040	-.02	-.021	**-.041**
L/H	**-.001**	.003	.004	.006	.007

Table 4 and Table 5 demonstrate that nuke attacks are also not effective and not able to achieve notable prediction shift values for varying filler size and attack size. Also note that love/hate attack is completely ineffective so that it cannot achieve even minus prediction shifts for most of the trials. Reverse bandwagon attack, on the other hand, is more effective than love/hate attack but obtained prediction shifts are negligible compared to the rating range. Therefore, we can again conclude that bisecting k-means clustering-based prediction algorithm is resistant against nuke shilling attacks in privacy-preserving environments.

4.2.3 Comparison with a memory-based approach

Robustness against profile injection attacks has not been investigated deeply in terms of privacy-preserving CF algorithms. A former analysis is performed to scrutinize robustness of memory-based k-nn privacy-preserving CF algorithm [15]. Therefore, we compare our experimental findings with such algorithm's perfor-

mance to provide a clearer ground for proposing bisecting k-means clustering-based privacy-preserving recommendation algorithm as a robust approach. Table 6 demonstrates maximum obtained prediction shift values for k-nn and bisecting k-means clustering-based privacy-preserving recommendation algorithms.

Table 6. Prediction shift values for k-nn and bisecting k-means clustering (BKM) algorithms

	Attack type					
	RN	AV	BW	SG	RBW	LH
k-nn	1.34	0.54	1.37	1.50	-1.75	-0.16
BKM	0.08	0.01	0.04	0.21	-0.05	-0.00

As can be followed in Table 6, bisecting k-means clustering-based algorithm performs highly robust and is exposed to insignificant prediction shifts compared to memory-based k-nn algorithm for all attack types. This fact originates from bisecting k-means clustering-based algorithm's success in separating most malicious profiles from genuine ones by recursively clustering them. Unlike bisecting k-means algorithm, k-nn approach is more open to shilling attacks as it directly calculates similarities with all existing profiles and set k nearest of them as neighbors.

5 Conclusions and Future Work

We investigated a formerly proposed private, accurate, and scalable bisecting k-means clustering-based prediction algorithm's robustness against malicious profile injection attacks. We presented modified versions of four well-known push and two nuke attacks to be implemented on privately collected databases. We also explained how such inserted attack profiles can affect the recommendation scheme and why it is expected that the algorithm is robust against them. According to the obtained experimental results, the demonstrated push and nuke attack models are not able to significantly alter final predictions produced by the scheme. Thus, the algorithm is robust against shilling attacks.

We are planning to investigate the robustness of the algorithm in non-private CF environment. It is known that clustering algorithms can be effective as a detection mechanism for shilling attacks. Hence, it warrants future work to utilize this algorithm as a detection tool of fake profiles.

Acknowledgment

This work was supported by the Grant 111E218 from TUBITAK.

References

[1] Alper Bilge, Serhan Gurmeric, and Huseyin Polat. An enhanced collaborative filtering scheme via recursive clustering. In *Proceedings of the Workshop on Knowledge Discovery, Data Mining and Machine Learning*, Dortmund, Germany, 2012.

[2] Olga Georgiou and Nicolas Tsapatsoulis. Improving the scalability of recommender systems by clustering using genetic algorithms. *Lecture Notes in Computer Science*, 6352:442–449, 2010.

[3] Manolis G. Vozalis and Konstantinos G. Margaritis. Using SVD and demographic data for the enhancement of generalized collaborative filtering. *Information Sciences*, 177(15):3017–3037, 2007.

[4] Michael J. Pazzani. A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, 13(5-6):393–408, 1999.

[5] Alper Bilge and Huseyin Polat. A comparison of clustering-based privacy-preserving collaborative filtering schemes. *Applied Soft Computing*, 13(7):2478–2489, 2013.

[6] Bhaskar Mehta and Thomas Hofmann. A survey of attack-resistant collaborative filtering algorithms. *IEEE Data Engineering Bulletin*, 31(2):14–22, 2008.

[7] Bamshad Mobasher, Robin D. Burke, Runa Bhaumik, and Chad A. Williams. Effective attack models for shilling item-based collaborative filtering systems. In *Proceedings of the 2005 WebKDD Workshop*, Chicago, IL, USA, 2005.

[8] Juan Lang, Matt Spear, and Shyhtsun F. Wu. Social manipulation of online recommender systems. *Lecture Notes in Computer Science*, 6430:125–139, 2010.

[9] Runa Bhaumik, Bamshad Mobasher, and Robin D. Burke. A clustering approach to unsupervised attack detection in collaborative recommender systems. In *Proceedings of the 7th IEEE International Conference on Data Mining*, pages 181–187, Las Vegas, NV, USA, 2011.

[10] Mark S. Ackerman, Lorrie F. Cranor, and Joseph Reagle. Privacy in e-commerce: Examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM Conference on Electronic Commerce*, pages 1–8, Denver, CO, USA, 1999.

[11] John Canny. Collaborative filtering with privacy via factor analysis. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 238–245, Tampere, Finland, 2002.

[12] Huseyin Polat and Wenliang Du. Privacy-preserving collaborative filtering. *International Journal of Electronic Commerce*, 9(4):9–35, 2005.

[13] Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. The impact of data obfuscation on the accuracy of collaborative filtering. *Expert Systems with Applications*, 39(5):5033–5042, 2012.

[14] Ihsan Gunes, Alper Bilge, Cihan Kaleli, and Huseyin Polat. Shilling attacks against privacy-preserving collaborative filtering. *Journal of Advanced Management Science*, 1(1):54–60, 2013.

[15] Ihsan Gunes, Alper Bilge, and Huseyin Polat. Shilling attacks against memory-based privacy-preserving recommendation algorithms. *KSII Transactions on Internet and Information Systems*, 7(5):1272–1290, 2013.

[16] David Goldberg, David Nichols, Brian M. Oki, and Douglas Terry. Using collaborative filtering to weave an information Tapestry. *Communications of the ACM*, 35(12):61–70, 1992.

[17] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1):5–53, 2004.

[18] Shlomo Berkovsky, Nikita Borisov, Yaniv Eytani, Tsvi Kuflik, and Francesco Ricci. Examining users' attitude towards privacy preserving collaborative filtering. In *Proceedings of the Workshop on Knowledge Discovery for Ubiquitous User Modeling*, 2007.

[19] Alper Bilge and Huseyin Polat. An improved privacy-preserving DWT-based collaborative filtering scheme. *Expert Systems with Applications*, 39(3):3841–3854, 2012.

[20] Chrysanthos Dellarocas. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of the 2nd ACM conference on Electronic Commerce*, pages 150–157, Minneapolis, MN, USA, 2000.

[21] Michael P. O'Mahony, Neil J. Hurley, and Guenole C. M. Silvestre. Collaborative filtering-safe and sound? *Lecture Notes in Computer Science*, 2871:506–510, 2003.

[22] Shyong K. Lam and John T. Riedl. Shilling recommender systems for fun and profit. In *Proceedings of the 13th International Conference on World Wide Web*, pages 393–402, New York, NY, USA, 2004.

[23] Shyong K. Lam and John T. Riedl. Privacy, shilling, and the value of information in recommender systems. In *Proceedings of the User Modeling Workshop on Privacy-Enhanced Personalization*, pages 85–92, Edinburgh, UK, 2005.

[24] Bamshad Mobasher, Robin D. Burke, Chad A. Williams, and Runa Bhaumik. Analysis and detection of segment-focused attacks against collaborative recommendation. *Lecture Notes in Computer Science*, 4198:96–118, 2006.

[25] Bamshad Mobasher, Robin D. Burke, Runa Bhaumik, and Chad A. Williams. Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. *ACM Transactions on Internet Technology*, 7(4), 2007.

[26] Ihsan Gunes, Cihan Kaleli, Alper Bilge, and Huseyin Polat. Shilling attacks against recommender systems: A comprehensive survey. *Artificial Intelligence Review*, pages 1–33, 2012.

[27] Alper Bilge and Huseyin Polat. A scalable privacy-preserving recommendation scheme via bisecting k-means clustering. *Information Processing & Management*, 49(4):912–927, 2013.

A Review on Data Mining Approach for Adverse Drug Reaction Research

Siriwon Taewijit[†‡] **Tu Bao Ho**[†] **Thanaruk Theeramunkong**[‡]
[†]Japan Advanced Institute of Science and Technology, Japan
[‡]Sirindhorn International Institute of Technology, Thammasat University, Thailand
[†]{siriwont,bao}@jaist.ac.jp,[‡]{thanaruk}@siit.tu.ac.th

Abstract

Tremendous researches in the drug development process currently play a crucial role on the post-marketing drug safety. Recently fast development of computational methods for adverse drug reactions (ADRs) has attracted high interest in the community of information technology. This work provides a survey on a vast variety of computational methods on ADRs. In particular, we mainly discuss sources of data and their characteristics, as well as a classification of computational methods based on data and techniques used, including their advantages and limitations. The challenges and further opportunities in this area are also provided.

Keywords: Adverse Drug Reactions (ADRs), Biological Network, Computational Network, Machine Learning, Data Mining, Text Mining.

1 Introduction

With significant rising of toxicology and clinical safety failures, drug safety and post-marketing surveillance have become a crucial topic [1], [2]. An adverse drug reaction describes harm associated with the use of given medications at a normal dosage during normal use, even a single dose or prolonged administration of a drug or result from the combination of two or more drugs. The number of deaths and reported serious from adverse drug events (ADEs) have rapidly risen from 5,519 to 15,107 and 34,966 to 89,842 respectively during 1998 to 2005 [3]. The World Health Organization (WHO) gives some key facts that motivate ADRs research as follows: ADRs are among leading causes of deaths and effect to in every country, the majority of ADRs are preventable, ADRs probably related to costs and no medicine is risk free [4].

The data-driven approach to ADRs where computational methods are applied to exploit clinical, omics, text and social network data is currently a wide subject in ADRs. Recently, Mckinsey Global Institute promotes a healthcare as one of the most challenging domains associated to big data to examine in depth and gain valuable contributions to global economy [5]. Drug safety is one leading research that plays a vital role on the drug development process. Almost 30% of developing drugs fail in a trial clinical stage due to toxicity [6] and estimated 2.5 million to 4 million of severe, disabling, or fatal adverse effects makes American suffer in the post-marketing stage [2].

Networks analysis such as biological networks, protein interaction networks, and medical social networks are well-known methodologies in the drug discovery [7], [8], [9]. The drug-protein interaction networks, which are construction of various kinds of networks integration, are immensely useful for ADRs prediction [10]. With rapid growth of online social networks and Web 2.0 technology, a large number of voluntary patients has led them to share their experiences on drugs, including usage and their side effects [11], [12] such online communities become valuable resources for ADR analysis. Moreover, in the past there were numerous effective machine learning-based and statistical approaches that aimed to construct a predictive model to reveal relations between drugs and side effects. Unfortunately most of them did not facilitate visualization or graphical representation of prediction results.

At present, study on ADR is still in its infancy stage, even it has high impact in the field of life science and human society. To facilitate researchers in this area, it is worth surveying the ADR-related data sources and potential computational models. Due to the limit of the length, this paper only used a very selected number of publications from a large ADR literature to examine existing drug side effect studies and discuss on perspectives of ADR opportunities and challanges.

Table 1. Summary of data sources

Data source type	Source of data	URL
Omics data	ChEMBL	www.ebi.ac.uk/chembldb
	ChemSpider	www.chemspider.com
	DrugBank	www.drugbank.ca
	GO	www.geneontology.org
	KEGG	www.genome.jp/kegg/kegg2.html
	PDB	www.rcsb.org/pdb/home/home.do
	PubChem	pubchem.ncbi.nlm.nih.gov
	SIDER	sideeffects.embl.de
	STRING	string-db.org
	SuperDrug	bioinf.charite.de/superdrug
	UniProt/ SwissProt	web.expasy.org
Clinical data	EMR/ EHR/ EPR	Hospital, Clinics, Physician offices, etc.
Text/Social networks data	DailyStrength	www.dailystrength.org
	DataGeno	www.datagenno.com
	PatientsLikeMe	www.patientslikeme.com
	Spontaneous reporting system	www.fda.gov/Drugs/GuidanceCompliance RegulatoryInformation/Surveillance /AdverseDrugEffects/ucm115894.htm
	WebMD	www.webmd.com

2 Data sources

Many computational methods aim to explore a relationship on ADRs or develop high performance ADRs predictive models. However, each of them can only work for one type of data that would be the first concern in ADR study. In general, we can widely identify three groups of data sources according to their characteristics.

2.1 Omics data

As large collections from various areas in biology, omics data provide a large pile of meaningful descriptions related to side effect. Some major data can be enumerated as follows.

Proteomics can be defined as a complete set of proteins produced by a given cell or organism under a defined set of conditions [13] while genomics is a studying about the genomes of organisms, e.g., genome sequencing, genome-wide association, gene expression analysis. UnitProt, SwissProt, PDB, GO, KEGG are the sources of proteomics data.

Metabolomics is the comprehensive, the qualitative, and the quantitative analysis of all the small molecules [14] or all chemical reactions involved in maintaining the living state of the cells and the organisms [15]. The related databases are such as ChEMBL, PubChem, DrugBank, ChemSpider etc.

Pharmacogenomics focuses on how our genes and complex gene systems influence our response to drugs [16]. The integration of KEGG, GO, DrugBank can provide information on this area.

Toxicogenomics is a scientific field and an outgrowth of the human genome project. It is closely allied to pharmacogenetics, which analyzes effects of heredity on responses of humans to drugs [17]. Furthermore, it describes the measurement of global gene expression changes in biological samples exposed to toxicants [18]. Several databases of this type are KEGG, UnitProt, SwissProt, STRING, DrugBank, SIDER.

2.2 Clinical data

There are three main types of clinical data ususally use in ADR analysis. Firstly, Electronic Medical Records (EMR) refers to a sort of patient medical information created in a hospital or an individual system as a digital version of a paper chart [19]. They include the outcomes from clinical laboratory testing, readmission summaries, patient medical records from one practice, diagnosis records, treatment plans and medication histories, etc. The second type, Electronic Health Records (EHR) is an integration of various different EMR systems, which created by medical institutes such as hospital, clinics, or physician offices. As privacy-protected and sensitive data, a traditional medical records consist of patient demographics,

medical histories, medications and allergies, immunization status, personal stats such as age and weight, and billing information etc. Lastly, Electronic Patient Records (EPR) can be described as electronic health records of individual patient (patient centric), provide primarily by one institute.

2.3 Online web, social media and text data

Recently, there have been several social media medical information sources on both web and mobile applications. DailyStrength [20] is a beneficial social network community on a healthcare domain. It is a wide diversity of medical, psychological and life conditions to discuss their struggles and treatments that they are undergoing.

PatientsLikeMe [21] provides a web based patients network. Herein, patients can share an encountered personal experiences with a disease in an online community. This community provides the online-data sharing platform, clinical trial awareness, finding other patients like them and learning from each other, which is particularly valuable in scientific researches.

WebMD, web based community that provides valuable health information services, health news, supportive communities, medical reference content databases, medical imageries, graphics, and animations, user experience, interactive tools, live web events, and in-depth reference material about health subjects.

The Spontaneous Reporting System (SRS) is the process of reporting of all unsolicited reports of ADEs from healthcare professionals, consumers to the Food and Drug Administration (FDA) or any appropriate authority [22].

3 Methods

Roughly, we can classify a number of computational models into two groups: biology-based approach and computation-based approach. Table 2 provides the summary for each method.

3.1 Biology-based methods

As biology-based approach, the computational biology methods can roughly be classified into three categories according to their core biological characteristics in the data used.

3.1.1 Ligand based methods

As a common approach, the ligand-based methods [23], [24], [25] presumes that if two drugs have a similar chemical structure, they usually have similar drug side effect. The advantage is simply to infer the drug side effect of another drug by manipulate chemical properties with known the drug side effect. However, the chemical structure cannot provide any biological interpretations regarding the underlying mechanisms at a molecular interaction level [10]. Moreover, sometimes they are different chemical structures but trigger the same drug side effect.

3.1.2 Biological information based methods

As the second approach, the biological information-based [26], [27] takes into account genes and pathways to more profoundly understand the biological process between a drug and a protein target. The assumption is any two similar biological process usually have the equivalent drug side effect. The advantage is a clear understanding about the biological process insight the relationship between drugs and their side effects. Nevertheless, this method is difficult for experimentally discovering the association between drug-targeted proteins and side effects in large-scale datasets in a cost-effective and an efficient manner.

3.1.3 Docking based methods

The third approach is based on docking [28], [29], [30]. It requires the three-dimensional structure information of a protein target to compare the similarity score of docking the drug into protein binding pockets. The best docking score results the potential similar side effects. Those methods are powerful to predict ADRs, however, the unsolved lacking of known three-dimensional structure remains as a problem.

A further point of view is to classify the method based on computational methods. Recently, several machine learning techniques take advantages of biological and social networks to uncover either relationships or predict unintended ADRs. The following describes rough categories based on computational methods.

3.2 Computation-based methods

3.2.1 Correlation based methods

The main process of this method is to calculate similarity score between two objects by us-

Table 2. Summary of methods

Methods	Sub methods	Data characteristic	Description
Biology-based	Ligand based	Chemical properties, Fingerprints, SMILES	Similar chemical structure → similar drug side effect
	Biological information based	GO, Pathways	Similar biological process → similar drug side effect
	Docking based	Protein 3D structure	Similar docking → similar drug side effect
Computation-based	Correlation based methods	GO, Pathways, Chemical properties	Correlation based technique
	Biological network based	Drug-Target protein network, Drug-Drug network, Protein-Protein network, Disease-Disease gene network, Drug-Therapy network	Network based technique
	Social net and text based	EMR, EHR, EPR, SRS, web communities	Text Mining
	Other machine learning methods	Protein sequence, PPI, GO, Pathways, Chemical properties, Fingerprints, SMILES, EMR, EHR, EPR, SRS, web communities	Feature based machine learning techniques such as SVM, HMM, Naïve Bayes, decision tree, neural network, regression, k-nearest neighbor etc.

ing correlation techniques. The score can represent the relationship between drugs and their side effects. In the first step, the method calculates and ranks the correlation scores. Consequently, the potential score is derived to make a decision on the side effect given drug. Recently, the calculation of linear combinations of two sets of random variables has been proved to be helpful for discovering relationships between each other using canonical correlation analysis (CCA). In [23], the correlation matrix of drug co-occurrence in protein-binding and side effect profiles can discover the extraordinary relationship. This approach applied data from DrugBank, SIDER, KEGG and Gene Ontology using adaptive sparse canonical correlation analysis. The results showed that most of significant correlated set has the same biological pathways, event if their molecular functions are different. Its preceding work [10] proposed a novel framework to predict ADRs of a given drug by considering on canonical correlation including network-based diffusion. The work can improve prediction performance from CCA approach.

Furthermore, the new effective method to predict ADRs and therapeutics is similarity ensemble approach (SEA) [31]. This is motivated by two concepts; the proteins can be related by their pharmacology and the established drug by exploiting the biological networks can exploit the relationship of side effects or protein targets. [32] uses ligand-based information from ChEMBL to examine the similarity using SEA. The closest similarity score between two drugs implies that they may occupy a same side effect.

Recently, Lee *et al.* [26] reveals the relationship between drugs and their side effects using side effect resources from SIDER and biological process information from GO. Their experiments are done to prove an assumption that phenotype can be impacted by action between drugs and biological processes. This research proposed three steps of process: (1) enrichment score calculation (ES), which expressed a degree of significant of set S (top or bottom) of the entire ranked list L [27]; (2) significant calculation, which applies t-score calculation for each process significance; (3) threshold-based filtering, which uses to filter drug side effect relations at twenty percent threshold of frequency.

Accordingly, two essential networks, drug-biological process network and drug side effect network were merged to derive the multi-level side effect related biological processes network.

3.2.2 Biological network based methods

One of representation of drug target relationships is a network construction and visualization.

This approach reveals a complex network of interactions and provides an intuitive analysis via network visualization. The mixture omics data from DrugBank, OMIM, and PPIs can be used for improving discovery of relations among protein targets, drugs, and diseases [7], [33], [34].

In [7], the method constructs the bipartite drug-target network (DT network) based on the common component sharing, which originates from the drug network and the target-protein network (TP network). In the drug network, any two nodes of the drug are interconnected, if they have at least one common target protein. On the other hand, in TP network, two protein nodes are connected to each other, if and only if both of them are targeted by at least one common drug. The clustering coefficient is used to consider topological features of a network. The result of this work supports that many proteins can be targeted by more than one drug, and most of the drugs have a distinct chemical structure.

The PROMISCUOUS [9] is proposed to provide the encompass data source of drug-side effect, drug-target, and protein-protein relations. The structural similarity of drugs is applied to discover relations among proteins, side effects and drugs. Although it can provide efficient network visualization, drug-target interactions have to be validated and tested experimentally due to their complexity of interaction characteristic.

3.2.3 Social net and text based methods

The social network and social media become famous source of data nowadays, due to influence of Web 2.0 technology. Many healthcare communities emerge in social networks. This can provide, service or exchange healthcare knowledge among users, e.g. DataGeno, PatientsLikeMe, and WebMD. The growth of online social networking forums has lead to patients voluntarily sharing of their experience on drug use, making these forums as valuable resources for ADR analysis [34], [35], [36].

Recently, the information extraction from the healthcare forum, Medications.com, has been proposed [12]. The source of data is the precious resource of sharing experience about drugs use among voluntary patients. Firstly, the implementation of Information Retrieval module can retrieve messages posted from the crawl technology. Secondly, applying Text Processing module can obtain an appropriate data format. Thirdly, the Information Extraction is manipulated to perform the Named Entity Recognition using drugs name and ADRs as a dictionary. Finally, Hidden Markov Model (HMM) is carried out to find the relationship between drugs and their side effects in Relationship Extraction module.

3.2.4 Other machine learning methods

The heterogeneities of omics data can bring opportunities when they are integrated. In fact, computational methods and text mining take into account to analyze ADRs relationships and generate predictive models. Utilization of chemical structures and biological properties or phenotypic characteristics were proposed in [37], [38].

Herein, four data sources were used to create high dimensional feature vectors. The encoded fingerprints corresponding to chemical structures are obtained from PubChem. The biological properties, targeted proteins, transporters, enzymes, and protein pathways, were collected from DrugBank and KEGG. The well-known drug side effect database, SIDER, provided the phenotypic of drugs, such as treatment indicators and other side effects.

To assess the effectiveness of the predictive model, five well-known algorithms namely logistic regression, Naïve Bayes, k-nearest neighbor, random forest, and SVM were employed. The results suggested phenotypic data were the most beneficial to predict ADRs. Moreover, SVM outperformed the others algorithm. However, when combined phenotypic data with baseline biological and chemical structures, the predictive model achieved signification improvements [39].

Not only omics data contains valuable information but also free text data contains relevant medication for ADR analysis. The FDA's Adverse Event Reporting System is a major source of ADR analysis. It is a large collection of ADRs reports, which remains challenging. Text mining techniques are usually provided to deal with this problem. Combination of multi-level text mining with informative feature selection and machine learning approach have been proposed for vaccine adverse event mining [40]. Another novel algorithm to detect drug-drug interactions in ADEs using SRS is proposed in [41]. The SRS from FDA is used to build profiles of 171 novel drug interactions for eight clinically adverse events classification. On one hand, MED-

LINE, medical literature is promoted as the text corpus to extract drugs-side effects, gene targets, and indications for cancer [42].

In [40], a large number of various classifiers were considered for advert prediction, including Naïve Bayes, maximum entropy, decision tree, recursive partitioning classification trees, boosted tree (BT), weighted support vector machines (w-SVM), SVM for sparse data (s-SVM), stochastic boosting, multivariate adaptive regression splines, regularized discriminant analysis. However, to avoid imbalance data set issue, calculated weights were assigned to all classes for BT, GAM, and s-SVM. The empirical results show that rule-based classifier, BT, and w-SVM performed well in terms of recall. Moreover, the rule-based classifier performed very well in terms of average sensitivity and specificity.

4 Discussions

In this paper, we survey effectiveness of existing ADRs methods. Most of research works can be categorized by either primary characteristic of data sources or technologies used. Firstly, we categorize research works into three groups based on biological perspective, ligand-based approach, biological information based approach and docking-based approach. Although these methods utilize various data sources, they share a common framework. The similar chemical structure of drugs or biological processes or protein targets can indicate that they are a similar drug side effect.

Many researchers revealed the biological process and the three-dimensional protein docking strongly related to drug side effects. However, lacking three-dimensional structure of protein target is a major obstacle. On the other hand, there are identical side effects, but also different drug chemical structures. Thus, the study on the relationship on biological process or combination of the sources of biological process and chemical structure can solve ADRs prediction.

From the computational view point, we can improve the effectiveness of correlation based method by providing feature selection in the early stage. The remaining challenge of this method is how to define suitable similarity measure. The network construction and visualization are widely used in biological network based methods. The network of drug-protein target, drug-side effect, protein-protein interaction, drug-disease, or protein target-biological pathways, can disclose crucial information. Moreover, they can improve the comprehensive understanding to analyze ADRs effectively. Many questions such as how to integrate many kinds of data sources, how to reconstruct the network, how to discover meaningful patterns from networks, are extremely challenging. From not only biological or chemical data sources but also conversations in social media. This powerful source can integrate and explore for the relationship towards the drug discovery process. The challenge arises up when we work on the variety and complexity of data sources.

Web 2.0 technology can promote social networks as a source of treasure. We can analyze ADRs via conversations which always flow in the social network conveyor. In addition, text information which expressed on technical reports of biomedicine, also useful. Text mining technique is usually applied to extract most frequent words or hidden topics, and the outcomes can utilize for the drug side effect analysis. In the future Web 3.0 technology, most data will be investigated and analyzed with semantics. To improve efficiency of text mining methods, semantics technology can enable machines as well as people to understand the meaning separately from data and content of conversations or reports.

Another way is computational methods such as SVM, neural network, k-nearest neighbor, Naïve Bayes, maximum entropy, decision tree, hidden markov model or regression. From the previous discussion, many researches can reveal that SVM expresses the outperformance rather than the others algorithm, and when we integrate data sources from biological and chemical data, this can achieve the higher performance.

At the time of the survey, there are varieties in ADR analysis methods such as statistics, biological network, social network, even computational analysis. However, this study is in an early stage. Integration of various data sources is a major challenge in this area. An exploring efficient algorithm to discover the relationship among objects in a heterogeneity network is also very interesting. Furthermore, high complexity of multi drug side effect is an advanced level of drug side effect analysis, which remains the opportunities and the challenges.

5 Conclusions

This paper presents a survey of existing solution for ADR analysis. In the past, there were varieties on methods and data sources that can be used. Most previous method can be classified into correlation-based method, biological network-based method, machine learning-based method and more modern social network-based method. In the future, more online data will be constructed and it is possible to reuse such information for ADRs. One interesting information for ADRs is the level of doses. With clinical information, it is possible to analyse relations between dose level and ADRs.

Moreover, in the real situation, often time set of drugs is provided to resist a disease. However this multi drugs therapy can lead to multi drug side effect issue, which is more complex than the traditional single drug side effect. This topic is interesting, challenging and open opportunity to study. We plan to extend our survey on the relation between multi drugs and their side effects.

Acknowledgment

This work was supported by JSPS Kakenhi Grant Number 23300105. Authors thank for the research environment, which was supported by the Japan Advanced Institute of Science and Technology (JAIST), Japan, and the Sirindhorn International Institute of Technology (SIIT), Thammasat University, Thailand.

References

[1] Institute for Safe Medication Practices. Monitoring fda medwatch reports: Why reports of serious adverse drug events continue to grow. 2012.

[2] Institute for Safe Medication Practices. Monitoring fda medwatch reports: Anticoagulants the leading reported drug risk in 2011. 2012.

[3] T. J. Moore, M. R. Cohen, and C. D. Furberg. Serious adverse drug events reported to the food and drug administration, 1998-2005. *Arch Inter Med*, 167(16):1752–1759, 2007.

[4] WHO Media centre. Medicines: safety of medicines adverse drug reactions. Fact sheet N293, 2008.

[5] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. Big data: Then next frontier for innovation and competition and productivity. *The McKinsey Global Institute*, 2011.

[6] Kola I. and Landis J. Can the pharmaceutical industry reduce attrition rates? nature reviews. *Drug Discovery*, 3(8):711–716, 2004.

[7] M. A. Yildirum, K. I. Goh, M. E. Cusick, A.-Lszl Barabsi, and M. Vidal. Drug-target network. *Nature Biotechnology*, 25(10):1119–1126, 2007.

[8] S. Hasan, B. K. Bonde, N. S. Buchan, and M. D. Hall. Network analysis has diverse roles in drug discovery. *Drug Discovery*, 17(15-16):869–874, 2012.

[9] J. von Eichborn, M. S. Murgueitio, M. Dunkel, S. Koerner, P. E. Bourne, and R. Preissner. Promiscuous: a database for network-based drug-repositioning. *Nucleic Acids Res*, 39(Database issue):D1060D1066, 2011.

[10] S. Mizutani, E. Pauwels, V. Stoven, S. Goto, and Y. Yamanishi. Relating drug-protein interaction network with drug side effects. *Bioinformatics*, 28:522–528, 2012.

[11] J. J. Mao, A. Chung, A. Benton, S. Hill, L. Ungar, C. E. Leonard, S. Hennessy, and J. H. Holmes. Online discussion of drug side effects and discontinuation among breast cancer survivors. *Pharmacoepidemiology and drug safety*, 2013.

[12] H. Sampathkumar, B. Luo, and X. Chen. Mining adverse drug side-effects from medical forums. *Imaging and Systems Biology (HISB) 2012 IEEE Second International Conference*, page 150, 2012.

[13] R.Twyman. *Principles of Proteomics*. Garland Science, 2012.

[14] G. Barchet. A brief overview of metabolomics: What it means, how it is measured, and its utilization. *The Science Creative Quarterly*, 8, 2013.

[15] R. L. Blaylock. *Natural Strategies for Cancer Patients*. Kensington Publishing Corporation, 2003.

[16] R. B. Altman, D. Flockhart, and D. B. Goldstein. *Principles of Pharmacogenetics and Pharmacogenomics*. Cambridge University Press, 2012.

[17] W. W. Weber. Toxicogenomics: History and current applications. *ASM News*, 70(8), 2004.

[18] Orphanides G. Toxicogenomics: challenges and opportunities. *Toxicol Lett*, 140-141:145–148, 2003.

[19] J. H. Carter. *Electronic Medical Records: A Guide for Clinicians and Administrators*. ACP Press, 2001.

[20] DailyStrength Inc. Health support groups at dailystrength. 2007.

[21] S. Gupta and J. Riis. Patientslikeme: An online community of patients. *Harvard Business School Case*, 511(093), 2011.

[22] S. R. Ahmad. *Spontaneous reporting in the United States. Chapter 9*. Stroms Pharmacoepidemiology, 2007.

[23] N. Atias and R. Sharan. An algorithmic framework for predicting side effects of drugs. *Journal of computational biology*, 18(3):207–218, 2011.

[24] E. Pauwels, V. Stoven, and Y. Yamanishi. Predicting drug side-effect profiles: a chemical fragment-based approach. *BMC Bioinformatics*, 12:169, 2011.

[25] J. Scheiber, J.L. Jenkins, S. C. K. Sukuru, A. Bender, D.Mikhailov, M. Milik, K. Azzaoui, S. Whitebread, J. Hamon, L. Urban, M. Glick, and J. W. Davies. Mapping adverse drug reactions in chemical space. *J. Med. Chem.*, pages 3103–3107, 2009.

[26] S. Lee, K. H. Lee, M. Song, and D. Lee. Building the process-drug side effect network to discover the relationship between biological process and side effects. *Bioinformatics*, 12(Suppl 2):S2, 2010.

[27] A. Subramanian, P. Tamayo, B. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43):15545–15550, 2005.

[28] Chen Y. Z. and Ung C. Y. Prediction of potential toxicity and side effect protein targets of a small molecule by a ligand-protein inverse docking approach. *J Mol Graph Model*, 20:199–218, 2001.

[29] I. Wallach, N. Jaitly, and R. Lilien. A structure-based approach for mapping adverse drug reactions to the perturbation of underlying biological pathways. *PLoS One*, 5(8):e12063, 2010.

[30] L. Xie, J. Li, L. Xie, and P. E. Bourne. Drug discovery using chemical systems biology: Identification of the protein-ligand binding network to explain the side effects of cetp inhibitors. *PLoS One*, 5(5):e1000387, 2009.

[31] Keiser M.J., Roth B.L., Armbruster B.N., Ernsberger P., Irwin J.J., and Shoichet B.K. Relating protein pharmacology by ligand chemistry. 25(2):197–206, 2007.

[32] E. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, B. K. Shoichet, and L. Urban. Large-scale prediction and testing of drug activity on side-effect targets. *NATURE*, 486:361–368, 2012.

[33] Y. Yamanishi, E. Pauwels, and M. Kotera. Drug side-effect prediction based on the integration of chemical and biological spaces. *Journal of Chemical Information and Modeling*, 52:3284–3292, 2012.

[34] L. Brouwers, M. Iskar, G. Zeller, V. van Noort, and P. Bork. Network neighbors of drug targets contribute to drug side-effect similarity. *PLoS One*, 6(7):e22187, 2011.

[35] M. Yang, X. Wang, and M. Kiang. Identification of consumer adverse drug reaction messages on social media. *Pacific Asia Conference on Information Systems (PACIS)*, 2013.

[36] C. C. Yang, L. Jiang, H. Yang, and X. Tang. Detecting signals of adverse drug reactions from health consumer contributed content in social media. *Proceedings of ACM SIGKDD Workshop on Health Informatics*, 2012.

[37] M. Liu, Y. Wu, Y. Chen, J. Sun, Z. Zhao, X. Chen, M. E. Matheny, and H. Xu. Large-scale prediction of adverse drug reactions using chemical, biological, phenotypic properties of drugs. *J Am Med Inform Assoc*, 19:e28–35, 2012.

[38] F. Hammann, H. Gutmann, N. Vogt, C. Helma, and J. Drewe. Prediction of adverse drug reactions using decision tree modeling. *Clinical Pharmacology & Therapeutics*, 88:52–59, 2010.

[39] L. C. Huang, X. Wu, and J. Y. Chen. Predicting adverse side effects of drugs. *BMC Genomics*, 12(Suppl 5):S11, 2011.

[40] T. Botsis, M. D. Nguyen, E. J. Woo, M. Markatou, and R. Ball. Text mining for the vaccine adverse event reporting system: medical text classification using informative feature selection. *J Am Med Inform Assoc.*, 18:631–638, 2011.

[41] N. P. Tatonetti, G. H. Fernald, and R. B. Altman. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. *J Am Med Inform Assoc*, 19:79–85, 2012.

[42] R. Xu and Q. Q. Wang. Toward creation of a cancer drug toxicity knowledge base: automatically extracting cancer drugside effect relationships from the literature. *J Am Med Inform Assoc*, 2013.

Data Mining for Life Insurance Knowledge Extraction: A Survey

Nanthawadee Sucharittham† Thanaruk Theeramunkong†

Choochart Haruechaiyasak* Bao Tu Ho Dam Hieu Chi****

† School of Information, Computer, and Communication Technology,
Sirindhorn International Institute of Technology, Thammasat University, Thailand
* Speech and Audio Technology Laboratory (SPT),
National Electronics and Computer Technology Center, Thailand
** School of Knowledge Science, Japan Advanced Institute of Science and Technology, Japan
{nanthawadee,bao,dam}@jaist.ac.jp, thanaruk@siit.tu.ac.th,
choochart.haruechaiyasak@nectec.or.th

Abstract

Although life insurance is useful reducing effect of unexpected vital accidents by sharing risk among a group of insurance members. Most people usually do not realize its importance, its benefit, and also do not know which type of life insurance should be acquired due to variety and complexity in insurance types. One of the challenging tasks is to facilitate a new customer to find a suitable insurance type that matches his/her behavior and life-style pattern in order to gain the highest benefits and satisfactions. This paper provides a survey on data mining techniques applied to extract knowledge from life insurance databases. The survey covers data sources, mined knowledge and mining techniques. Finally some suggestions are given as a guideline for research on mining life insurance data.

Keywords: service science, intangible product, customer satisfaction, knowledge extraction

1 Introduction

Recently the service sector becomes larger and gains more impact in our society. However, according to Spohrer et al., the service sector accounts for most of the world's economic activity, but it is the least-studied part of the economy [25]. Some recent promotions of research on service and its system includes IBM's project named Service Sciences, Management and Engineering (SSME), Oracle and IBM's industry consortium called the Service Research and Innovation Initiative and several service companies' service-value networks, such as those in Google, Amazon, Airlines, UPS, Wal-Mart. More details can be found in [11].

As a classic service but complex structures, insurance becomes an interesting target field of study for service science. Currently due to various conditions and needs, a large number of insurance types and policies are developed. With a suitable type and policy, the customer and/or his/her family can sufficiently obtain income when an unexpected inconvenient event, such as accident, sickness or fatal death, takes place. Looking as a whole, good insurance plans can reduce risk of both customers and insurance companies. However, the life insurance is also not so popular in most developing countries including Thailand due to difficulty of insurance terminologies, complexity of insurance policies and intricacy of benefit assessment. Consulting proficiency is required to extract a customer's behavior and life-style from huge data in order to describe effect of each policy to the customer and to determine the level of coverage for an individual people [20].

Recently information technology has been developed to the level of practice. Techniques in knowledge discovery and data mining (KDD) are studied. Especially, purchase and claim patterns mined from insurance databases can be used to suggest current and perspective customers to reduce risk and maximize beneficial outcomes.

In this paper, we survey possible application of data mining techniques on life insurance data, including those related to group insurance, non-life insurance and life insurance. The purpose of this paper is to review past literatures, to discuss on data types and data mining provided for extracting interesting patterns or knowledge from data in insurance industry. The rest of this

paper is organized as follows. Section 2 describes insurance knowledge extraction. Challenging tasks in mining insurance data are explained in Section 3. In Section 4, a conclusion is made.

2 Insurance Knowledge Extraction

Knowledge utilization provides an insightful understanding for people and organization such as allowing an operation in accordance with their demands, making coordination between the parties, or facilitating the management of information to make better decisions. That means knowledge is valuable however it needs some actions to find the way to make use of knowledge on hand. Knowledge extraction is the process to explore and construct knowledge from various sources of data. This knowledge is extracted into the solid results for easy understanding and we can utilize this useful knowledge to support our purposes.

One of the valuable data sources in a company is its database storing the customer activity records. It is possible to analyze customer behavior from such database, including both structured and unstructured data. For instance, the structured data might consist of demographic information including private information. An unstructured data is narrative information such as opinions from customers and evaluation results of underwriting.

By utilizing knowledge extraction from these structured and unstructured data in the insurance industry to analyze customer behavior, we can understand the way to obtain, use and dispose products. A new customer can consume this knowledge acquired from the current customer behavior if there are the same behavior, life style and conditions.

However, it is not straightforward to figure out which insurance policy can be recommended to a specific customer since even if the most similar customer, he/she may occupy minor different characteristics, life-style and conditions. Challenging in this task is how to collect a large amount of data for analysis and how to analyze such big data.

Customer behavior can be exploited by using knowledge extraction to support in various purposes from insurance activities. The initial criteria which can determine the basic rules for assigning an insurance type to an individual customer have three viewpoints. In this paper, we present how data mining could be used to extract knowledge to reply for three initial criteria as follows.

2.1 Which group or segment of a customer likely to be? This answer shows a position of a customer on current market situation before making new contract from customer segmentation.

2.2 Which product would best suit the customer needs? This answer shows the patterns of product concurrence set which present the customers behavior when they bought the set of products.

2.3 What is the risk rating of customer related to the amount of assurance? This answer shows the prediction for risk level of each customer. Company can also predict the loss when accept this case to be insured.

2.1 Customer segmentation extraction

The important of customer segmentation is to make new customers know their current positions which segment matches up character or life style. Because the customer separations is used for separating the group by determining the similarity of member make more understand the special characteristics of each customer segment. When new customers know themselves or business could comprehend the outstanding characteristics of each customer segment, they can propose the right product or right service to the right segment. A traditional method to propose new product to new customer is to select by simply target from basic knowledge base or experience of selling-agents or another marketing channel. From this reason, it is necessary to identify the segment of each individual customer utilize some cluster analysis techniques to improve the traditional method.

Generally, insurance company gets demographic data of customer when customer made new contract. They can utilize demographic data of customer such as education, gender, income level, marital status, age, premium, and life stage for extracting customer segmentation. One of data mining technique called “clustering technique” which attempts to group a set of given data into classes, so that members within cluster have high similarity in comparison to another but these members are very dissimilar to other clus-

ters. Similarities or dissimilarities are assessed based on the attribute values describing the objects where distance measures are used. In generally, the clustering has many techniques such as partitioning method, hierarchical method, density-based method, K-means clustering. However we give an example how clustering technique works in the area of customer segmentation from K-means clustering.

Customer segmentation can perform extraction using a clustering techniques based on determined attribute values to create new catalogs. The expected result from this analysis target is to find out the position of new comers in the current insurance market. A new customer will know which segmentation matches to his character or his life style by demographic customer data analysis.

Refer to Table 1, there are provided attributes from history demographic customer data such as gender, age, occupation, income. An analysis attribute is life stage attribute. Customers who have attention to make insurance contract should know their position from the results after exercise clustering techniques.

Table 1: Sample history demographic customer data

Gender	Age	Occupation	Income	Life Stage
Male	7	Student	0	Juveniles
Male	21	Soldier	15,000	First Jobbers
Male	61	No job	15,000	Retirees
Female	14	Student	0	Juveniles
Female	35	Officer	35,000	Matured
Female	55	Teacher	65,000	Matured

The data from the determined attributes can perform a clustering to separate the cluster of customer. In brief, the steps of K-means consist of do partition objects into non-empty subsets, compute mean point of the clusters of the current partition, assign each cluster center to the mean of its assigned item and then repeat until convergence. Since clusters are not predefined, a domain expert is often required to interpret the meaning of the created clusters. The supposed results after took experiment from K-means clustering consists of a first cluster is male, 32, soldier, 16,500, first jobbers and a second cluster is male, 35, officer, 25,000, matures in terms of gender, age, occupation, income and life stage. Next, we put a test data as male, age is 30, occupation is officer, and income is 20,000. This case the cluster of this test data is the cluster one and life stage should be first jobbers. It might be different from find out result from basic analysis which may classify this case to be matured.

Many researchers often utilized clustering techniques in their tasks. Various papers used K-means algorithms in their experiments such as A.B. Devale and Dr. R.V.Kulkarni used K-means clustering technique for classify the group of customer from demographic information such as age, occupation, income and education level [2]. This paper introduced the basic concept of each data mining concept for knowledge discovery in insurance business. In clustering, they used K-Means in order to give an example in case of company does not have any predefined any labels. They presented based on the outcome of the grouping they will target marketing and advertising campaigns to the different group for particular types of policy. Shu-Hsien Liao et al. utilized the concept of data mining in real work to determine the level of customers' needs in life insurance products by extracting specific knowledge patterns and rules from consumers and their demand chains [27]. This paper also used K-means clustering to classify the customer group from demographic data which is quite more different features they not only used basic information as name, gender, birth date but also included life stage needs, residential house level, life partner, and transportation mode. Depth details of features can distinguish classification of customer groups more clearly.

Many previous works classified customer groups in different purposes from clustering method. Not only clustering method but also other methods can utilize for customer segmentation. For instance, the association rules are widely utilized especially market baskets theory for finding existing customer patterns such as Yu Yan and Haiying Xie recommended using association analysis in customer behavior model to improve the value of customer in customer relationship management task (CRM) of insurance industry [29]. For example, the recommendations of new insurance service to old customer in order to upgrade sales volume and improve return rate of sale-investment. Moreover, they also gave some recommendations about an important level of informatization in pattern construction. Yongqiang Chen & Leifang Hu employed data mining knowledge into customer relationship management (CRM) in insurance business to enhance the analysis of customer value model and market classification model [28]. In part of

customers values model constructed from purpose of lifetime value conception from formula (customers' lifetime value = potential value + current value) by adapt knowledge of economic into data mining process. In sum, the final model got potential value of insurance equal to customers has potential to purchase policy plus customers' recommendation value plus the other value. Market classification models were constructed by decision tree from huge number of customer data. Using data mining can distinguish different sorts of customer to be many customer classification as customer satisfaction, customer loyalty, customer declining, individualized service, etc.

2.2 Occurrence product patterns extraction

Finding out the suitable of life insurance product for each individual people is a difficult task because there are various kinds of insurance product which is not only basic product type but also rider concurrence product. They are very difficult to understand the useful of the different purpose of each product and difficult to know which group of product should be hold simultaneously in order to support the highest benefits.

At present, selling channels or agents still offer a new product to a customer from their experience without considering the highest benefit of customer because it is not straightforward to understand clearly for the relationship of product benefits between basic and rider so they cannot recommend the right product to the right person correctly. Normally, they offer base on their idea and sometimes base on their commission of each product at that time.

Normally, life insurance product has two product types. First type called basic product which means one main policy have only basic product for each insured. Second type called rider product which is supplement product into basic product. Each basic product has different purpose. Some of basic product has purpose to build up cash value and others is a protection purpose in short term. The example of basic product in different purposes such as [17]

- Term life insurance is for a limited time such as 10, 15, or 20 years.
- Universal life insurance is flexible and allows to increase premium over time
- Whole life insurance is permanent insurance and provides lifetime coverage. The premiums are fixed and builds cash value. The value functions as a saving account that may be tax deferred.

Rider product is the additional benefits that can be added into a basic product. Rider product is increase a level of coverage and can be blended, for an additional cost, according to future coverage needs [19]. Here is some examples of common riders in life insurance.

- Waiver of Premium (WP) is a rider attached to the basic policy to maintain coverage in the event of total and permanent disability; subsequent premiums are waived.
- Hospital & Surgical Benefit (H&S) rider provides reimbursement for medical expenses in case of hospitalization in a licensed hospital as an in-patient (IPD)
- Accident Death Benefit (ADB) is one-time payment after insured is diagnosed with a terminal illness that will considerable shorten lifespan.

Each insured hold the different kind of riders within one policy which are under only one basic product. That means it has different kind of riders that are often sold together. In order to solve this problem, it is necessary to search the pattern of product group which suit for each individual customer. Some techniques from data mining such as association rules can help extracting the hidden product concurrence group from previous history purchasing of customer data. The target results are finding all the association patterns from various features of history customer data that present the customers behavior when they bought products as a subset of frequent item sets, most of the time also bought the remaining items in the same frequent item set.

Many researchers used association rules mining technique such as apriori algorithm and market basket analysis which are techniques for generating rules from a large data set based on the co-occurrences among the items. The results from this analysis is to learn life insurance item sets of basic and rider product concurrence in order to understand what customer is demanding and these patterns can be used for recommend product sets. The concept of association rules is below.

Association rules mining results in the implication of the form X => Y (If X then Y), where X and Y are the itemsets

Step 1: finding the candidate itemsets which pass

the minimum support[1]; Support (X => Y) = P(X and Y)

Step 2: generating rules from itemsets which pass the minimum confidence[2] (accuracy)

Confidence (X=>Y) = P(Y|X) = P(X and Y)/P(X) = support (X => Y)/support(X)

Table 2: Sample product occurrences set

Item ID	Item sets			
	Basic	Rider		
T1	Term	WP	H&S1	H&S2
T2	UL	WP	ADB	
T3	WL	WP	H&S2	
T4	UL	WP	H&S1	ADB
T5	WL	WP	H&S1	H&S2

Remark: UL – universal life, WL – whole life, WP = waiver of premium, H&S1/2 = hospital & surgical benefit type 1/2, ADB = accident death benefit

Table 2 is a sample product occurrence data set which consists of basic and rider product in these insurance policies. We presents this example from produces the association rule generation with 3-itemsets rules under condition minimum support = 30% and minimum confidence = 80%. The results found that when people bought one basic product 100% bought one waiver of premium (WP) and at least one type of hospital & surgical benefit (H&S). In terms of rider occurrences, 100% when bought accident death benefits (ADB), people bought waiver of premium (WP) and hospital & surgical benefit (H&S).

Association rules mining plays an important rules to find out hidden pattern from basic product sell concurrence with rider product. Besides, many researchers also use a statistical concept to analysis this point such as K-nearest neighbor, pearson correlation. A.B. Devale and Dr. R.V. Kulkarni discussed how data mining method perform the sophisticated classification and correlation between policy designing and policy selection [2]. They presented association rules for determined the frequent item sets based on a predefined support like a riders that are often sold together with basic products. They used market basket analysis to determine capitalized on the association between different policies that are sold for different purposes. In part of classification algorithms, A.B. Devale et al. presented classification method of K-nearest neighbors which was used for classifying product selection in each age range and occupation, providing big pictures how to benefit from K-nearest neighbors by do segmentation based on income, paid premium, premium mode and net of assurance price. They showed many examples to classify the product type in relationship of each features above such as “Life security”, “Tax benefit”, “Investment” [2].

Nan-Chen Hsieh & Kuo-Chung Chu used K-means classified customers into five clusters and then applied apriori algorithm as association analysis for each product group in these customer clusters [18]. This experiment investigated what functionalities best fit the customer’s needs in life insurance product by extracting the specific knowledge patterns and rules from customer behavior and their demand chain. Generating new product needs not only customer data but also risk rate, ridership and ridership were modeled based on limited data. The main purpose is to share the risk and reduce premiums after adding in expense and profit, a final premium is determined more accuracy.

In order to improve the performance of product recommendation, we can utilize the benefit by integrating knowledge from customer model to product model to improve this process. Knowledge Discovery in Databases (KDD) process in data mining can help merging customer data from different platforms. A.B. Devale and Dr. R.V.Kulkarni presented the statistical principles can be enhanced the capabilities of data mining method to improve directly into insurance procedures [2]. For example, Pearson correlation of statistical methods to specific correlation coefficient in positive and negative correlation was used for finding out a correlation between policy designing and policy selection.

2.3 Risk assessment

Risk assessment is one of the core functions of insurance business. It estimates a likelihood and impact of risk in order to represent the different level of impact with this risk. Insurance industries estimate and set the risk rating for predict the loss from accidents, health claims, or disaster rates in order to support the underwriting process which set up risk rating of customer by placing customer in a predetermined risk class with ambiguous names such as “super preferred”, “preferred”, “standard” and “substandard”. This rating is estimated from customer health and physical characteristics (such as age, gender, height, weight, alcohol, drug, tobacco use or extreme sports) and this rating will impact directly to

[1] Support factor is an indicator that informs how much occurrence of the events X and Y.

[2] Confidence factor is an indicator that informs how many percentage of Y happen when X happen.

premium rate which customer must pay to make contract with insurance company. Traditional method made use of actuarial statistics within the difficulty to predict base on stochastic in nature. It is necessary to construct risk prediction model from some techniques. Data mining and various fields of machine learning, statistics, and optimization techniques are utilized.

In this paper, we give some examples of previous works from various kinds of techniques. The specific analytical techniques from data mining such as market basket analysis, clustering technique, and statistical method can help improve the fraud claim process. IBM T.J.Watson research center [4, 5] also developed method called "*ProbE*" (probabilistic estimation) predictive-modeling algorithm. This method uses *C*++ kernel *rule* based models of insurance risk, where each rule represents a risk group. The purpose of this method is for underwriting profitability analysis which mines property and casualty insurance policy and claims data to construct predictive models for insurance risk. In actuarial science, accuracy is measured in terms of statistical confidence intervals that are estimated risk parameters can deviate from true values. In this experiment, there are concerning point about avoiding over-fitting, it occurs when the best model relative to the training data tends to perform significantly worse when applied to new data.

Jianbing Xiahou and Yang Mu presented the comparable result between traditional approach and data mining approach [15]. Traditional method is statistical using generalized linear models (GLMs). It is an extension of general linear function $Y = X^T\beta + \varepsilon$ which consists of three factors: random factor means Y follows a normal distribution, $\mu = EY$; system factor, namely $\eta = X^T\beta$ (X^T is transport matrix of X); connection factor, namely h=m. In practice, Y does not follow a normal distribution. Researchers also comment in the general statistic method in this experiment. This task needs to process data several times to fitting in practice since the section of risk variables and classification of variables are not very exact. That means the cost and complexity are huge because input matrix is huge. Moreover, the models they get by fitting may not be convergent. They also compared results of experiment to select the risk in classification using decision tree algorithm to rate making into risk analysis. They expected decision tree classifier can solve the problem of multitudinous information classification. However the results form this experiment still did not satisfy since they found risk variables which make an important impact on rate making are too few in their models. Besides, the risky variables provided by insurance agents are not ranked reasonably. Yu Yan and Haiying Xie explained the useful of data mining in task of dividing credit customers into several grades, predicting the customer risk, making credit rating and avoiding service risk [29]. They also explained in part of classification and prediction to construct forecasting model of future tendency of data using decision tree, neural network, and Bayesian network.

3 Data mining challenging for knowledge extraction in insurance business

An efficiency of knowledge extraction is depend on a quality of data which need all preprocessing processes in KDD especially cleansing data to handle missing data, reducing noise and accounting for time series. However, we still face with some obstacles of high volume and complexity of data and now many researchers are interested in solving problem with data mining hybrid techniques to get high performance results.

3.1 High volume and high complexity in the different kinds of data

There are the different kinds of data types used in data mining process such as structural and unstructured data type. A structured data is readily searchable by simple, straightforward search engine algorithms; whereas an unstructured data is essentially the opposite. All data types are collected in different repositories such as relational databases, time-series database, data streams or multimedia database. We need to integrate all data platforms into the same platform before extract knowledge by data mining technique. Most of historical data in company is the structured data however the data has huge volume of data stream. The unstructured data also plays important roles now because some data in company is narrative platform and there are plenty of free texts such as customer comments are retrieved from various websites. The unstructured data needs some necessary steps to

transform unstructured data to be a structured data before perform data mining by some techniques such as natural language processing (NLP) and human language technology (HLT).

Pennock et al. performed a semantic classification of reviewing messages in unstructured data using natural language processing by N-gram and be continued data mining processes to compare the results between support vector machine (SVM) and Naïve Bayes from C|Net text corpus [6]. Hu & Liu extracted the co-occurring sets of terms from unstructured data then they did feature generation from these occurrence sets and used association rule mining to find all frequent items set. In this context, an item set is a set of words or phrases that occurs together. They used apriori algorithm in step of find all frequent item sets by satisfy a user-specified minimum support however they did not construct any sequent patterns. In sum, the process of unstructured data utilization needs more complicated methods such as NLP, IR to preparing data before extract knowledge from data mining algorithms [3, 12, 13].

3.2 Hybrid techniques in data mining research

In order to make data mining method more powerful, many researchers presented novel methods by consolidated data mining algorithms mixed together or consolidated with other knowledge such as artificial intelligence, information retrieval (IR), NLP. For example, Nan-Chen Hieh & Kuo-Chung Chu presented in the topic of enhancing customer behavior analysis by data mining techniques [18]. This paper showed two-stage frame work of consumer behavior analysis, and the key feature is a cascade involving self-organizing map (SOM), a neural network and a decision tree inducer. This hybrid framework used clustering techniques to preprocess input samples into homogeneous clusters, and decision tree techniques to build customer profile. This paper did the clustering customer behavior from general demographic information such as repayment cycle, number of purchases, card age month, block code, sex and credit line. They gave some suggestions about the clustering algorithm affects the performance of the clustering result. The quality of input samples leads to misclassification

Shu-Hsien Liao et al. investigated what functionalities best fit the consumers' needs in life insurance products by extracting specific knowledge patterns and rules from consumers and their demand chains [27]. This paper used apriori clustering algorithm to illustrate the marketing segmentations and demand chain on life insurance in Taiwan. They also did segmenting of potential insurance buyers into similar groups using K-means and association rule based on buyer's previous purchased.

4 Conclusion

We investigated the knowledge extraction from data mining techniques used for life insurance. Main purpose is to provide more understanding in the current customer behaviors and to utilize them for recommending a suitable policy to new customers in order to gain their highest benefits and satisfactions. Company also has an efficient customer relationship management and gain higher persistency rate. Various data mining methods were used for knowledge extraction. Each method has difference strengthen points to utilize them. In this paper, we gave some examples of data mining methods to give some ideas to utilize them.

Furthermore, we found that many researchers are interested in hybrid techniques to make knowledge construction more efficiency and the correctness of knowledge extraction from data mining need to overcome the obstacle from high volume and high complexity from the different kind of data such as structured and unstructured data which generated from diverse medium.

Acknowledgment
This research was conducted under a grant in the SIIT-JAIST-NECTEC Dual Doctoral Degree Program.

References

[1] Abhijit Raorane and R.V.Kulkarni, Data mining techniques: A source for consumer behavior analysis, International Journal of Database Management Systems,Vol.3,No.3, August 2011.
[2] A.B. Devale and Dr. R.V.Kulkarni, Application of data mining techniques in life insurance, International Journal of data mining&knowledge management process Vol.2, No.4, July 2012.
[3] Bing Liu, Ding, Xiaowen, and Philip S. Yu. A holistic lexicon-based approach to opinion, WSDM'08 Proc. of the International Conference

on web search and mining, 231-240, 2008.
[4] C.Apte, E.Grossman, E.Pednault, B.Rosen, F.Tipu, and B.White, Insurance risk modeling using data mining technology, IBM T.J. Watson, RC 21314 (94531) March 31,1998.
[5] Chidanand Apte, Edna Grossman, Edwin P.D.Pednault, Barry K.Rosen, Fatch A. Tipu, and Brian White, Probabilistic estimation-based data mining for discovering insurance risk, IBM T.J. Watson research center, IBM Research Report RC-21483, IEEE 1094-7167/99,1999.
[6] Dave, Kushal, Steve Lawrence and David M. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. Proc. of the 12th international conference on World Wide Web, 519–528, 2003.
[7] Dennis Wegener and Stefan Ruping, On Integrating Data Mining into Business Processes., In Proceedings of PADD99,2010.
[8] Dr Usha Thakur, ATS research and Vinay Kumar Patri, BFSI solutions, HCL technology of Chennai, CRM Trend in insurance, HCL Technologies Ltd. April 2010.
[9] Eui-Hong (Sam) Han, George Karypis and Vipin Kumar, Min-Apriori:An algorithm for finding association rules in data with continues attributes, Expert systems with application39,11303- 11311, 2012.
[10] E.W.T Ngai, Li Xiu, D.C.K. Chau, Application of data mining techniques in customer relationship management: A literature review and classification, Expert systems application 36, 2592 – 2602, 2009.
[11] Hidaka, K., Trends in services sciences in Japan and abroad. Quarterly Review 19 pp. 35-47, April 2006.
[12] Hu, Minqing and Bing Liu, Mining Opinion Features in Customer Reviews., In Proc. Of American Association for Artificial Intelligence(AAAI), www.aaai.org, 2004.
[13] Hu, Minqing and Bing Liu, Mining summarizing customer reviews. Proc. of the SIGKDD international conference on Knowledge discovery and data mining, 168–177, 2004.
[14] Investopedia, Life Insurance, www.investopedia.com/terms/l/lifeinsurance.asp
[15] Jianbing Xiahou and Yang Mu, Analysis on application of data mining upon rate making and risk analysis in insurance, Computer Engineering and Technology (ICCET), 2010 2nd International Conference on (Volume:2), April 2010.
[16] Jim Spohrer and Paul P. Maglio. The emergence of Service Science: Toward systematic service innovations to accelerate co-creation of value, IBM Almaden research center, Production and Operations Management Society, Vol. 17, No. 3, 238–246, May–June 2008.
[17] Micah, A Basic Guide to Choosing a Life Insurance Plan, Feb 20, 2012, www.money-choice.org/basic-life-insurance-plan-guide
[18] Nan-Chen Hsieh and Kuo-Chung Chu, Enhancing customer behavior analysis by data mining techniques, International Journal of information and management science, 39-53, 2009.
[19] Pooja Dave, Let Life Insurance Riders Drive your Coverage, Feb 26, 2009, www.investopia.com/articles/pf/07/life_insurance_rider.asp
[20] PR Newswire Association LLC, Deloitte, Consumers Eager for Life Insurance Survey Says, May22,2012. http://insurancenewsnet.com/oarticle/2012/05
[21] Preeti Sharma and Sanjay Kumar. An Approach for Customer Behavior Analysis using Web mining. International Journal of Internet Computing, ISSN: 2231-696, Issue-2, 2011.
[22] Sanford Gayle, SAS Enterprise Miner Software, A SAS institute white paper: data mining in the insurance industry, 1999.
[23] Se June Hong and Sholom M. Weiss, Advance in predictive models for data mining, Pattern Recognition Letters 22, 51 – 61, 2001.
[24] Sitikantha Kar and FerozKhan A L, Mining Unstructured Text Data for Insurance Analytics, Tata Consultancy Services, August 2010.
[25] Spohrer, J., Maglio, P., Bailey, J. and Gruhl, D., Steps toward a science of service systems. IEEE Computer Magazine 71-77, January 2007.
[26] Shu-Hsien Liao, Pei-Hui Chu and Pei-Yuan Hsiao, Data mining techniques and applications – A decade review from 2000 to 2011, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Vol:39, Issue: 2, March 2009.
[27] Shu-Hsien Liao, Ya-Ning Chen and Yu-Yia Tseng, Mining demand chain knowledge of life insurance market for new product development, Expert Systems with Applications: An International Journal archive Volume 36 Issue 5, 9422-9437, July 2009.
[28] Yongqiang Chen and Leifang Hu, Study on Data Mining Application in CRM System Based on Insurance Trade, Proceedings of the 7th international conference on E-commerce, 2005.
[29] Yu Yan and Haiying Xie, Research on the application of data mining technology in insurance informatization. Ninth International Conference on Hybrid Intelligent Systems, 2009.

Classification of Mammogram Images using Support Vector Machine

Karmilasari† Suryarini Widodo‡* Lussiana ETP* Matrissya Hermita Lulu Mawadah ***

† ‡* ***Faculty of Computer Science and Information Technology, Gunadarma University, Indonesia
*Study Program of Information System, School of STMIK Jakarta STI&K, Indonesia
**Faculty of Psychology,Gunadarma University, Indonesia
{karmila,srini,matrissya}@gunadarma.ac.id
lussiana@jak-stik.ac.id
lulu_chester91@student.gunadarma.ac.id

Abstract

Breast cancer is one of the main causes of cancer death in women. Early detection is eficiently performed by using Digital mammograms. This research will be developed methods that can be used to detect abnormality of breast cancer. Step by step to classify normal and abnormal classes on digital mammogram is image acquisition which images are taken from MIAS database, find the region of interest using morphology algorithm, feature extraction using GLCM (Gray Level Cooccurrence matrix) and the last is classification process using SVM (Support Vector Machine). The accuracy of abnormalities (normal or abnormal) on digital mammogram for each testing data used in classification process have a different value, that is 85.00% percentage were taken inside from training data, 60.00% were taken outside from training data and 67.5% were taken inside and outside from training data. From these results, testing data that used in classification process affect on percentage of abnormality accuracy. For the future, is expected in this system can determine the level of severity (benign and malignant) from classification results.

Keywords: Mammogram, GLCM, SVM

1. Introduction

Breast cancer is by far the most frequent cancer among women with an estimated 1.38 million new cancer cases diagnosed worldwide in 2008 (23% of all cancers), the number of deaths by 458 and ranks second overall (10.9% of all cancers) [3]. Indonesia Health Profile 2008, published by the Ministry of Health Republic Indonesia display the data in 2004 to 2007 the number of breast cancer cases occupies the top position, followed by cervical cancer [7].

The early symptoms of breast cancer is often not recognized or perceived by the patient. Thus, many patients who seek treatment in an already severe. Breast cancer is characterized by a lump in the breast. The lumps can be benign or malignant tumors. One way should be done by women to avoid breast cancer is early detection, such as breast self-examination.

Technological developments, especially in the field of medical imaging have made the detection of cancers such as breast cancer. In which the cancer can be more easily detected, one of them with mammogram. A mammogram is a tool that can generate two-dimensional, usually 8-bit gray scale image that obtained from the x-ray of the breast patients [10]. Visually, a doctor can identify breast abnormalities by looking at the characteristics of the mammogram image. These characteristics are the left and right breasts are not symmetrical, there is a lump, there is spread of breast tissue structures and contained microcalcifications [12].

From mammogram image, the doctor performs the conventional analysis or directly diagnose cancer cells with the naked eye [10]. However, sometimes there is a mistake when analyzing (due to fatigue and various kinds of human error) as well as let loose the important things from eyesight doctor.

Based on this background, it will be

developed methods that can be used to detect the presence of breast cancer. Some of the methods developed to detect the presence of breast cancer are recognizing abnormalities on mammography images and detection of breast tumors. Step by step to classify breast cancer is image acquisition, segmentation process or find the region of interest, feature extraction, and finally, classification process.

According to Liu, structural abnormalities in mammography images can be recognized through the existence of microcalcifications, a lump and distribution lesion or nodule. Normal characteristics are shown in mammogram image that have a lower density than abnormal mammography image, there are no stains that showed micro calcifications and there is no brighter areas (higher intensity), which showed a lump [11].

To diagnose breast cancer on mammography images, then performed region of interest selection process or the suspected area. In the research Basim Alhadidi, Mohammad H.Zu'bi and Hussam N.Suleiman, breast cancer can detect from mammogram image using image processing functions : threshold, edge based and watershed segmentation. From these method shows that threshold is the fastest. But the output image from threshold is not segmented clearly than other method, edge-base and watershed segmentation consuming long time than threshold but output image is better, also threshold only segment gray level images [5].

After found the region of interest, then do extraction features. In the research Pradeep N, Girisha H., Sreepathi B. and Karibasappa K., develop a computer aided method for mass or tumor classification based on extracted features from the Region Of Interest (ROI) in mammograms. ROI has to be segmented from the digital mammogram using the Segmentation techniques. Pattern recognition in image processing requires the extraction of features from regions of the image, and the processing of these features with a pattern recognition algorithm. They consider the feature extraction part of this processing, with a focus on the problem of tumor detection in digital mammography. Parameter used consists of 12 features [13].

Y.Ireaneus Anna Rejani, Dr.S.Thamarai Selvi, proposed system focuses on the solution of two problems. One is how to detect tumors as suspicious regions with a very weak contrast to their background and another is how to extract features which categorize tumors [15].

Based on the previous research, it will created a system that can assist doctors in the diagnosis of breast cancer. With these system, is expected to classify the mammogram image into the class of normal and abnormal (malignant or benign). So it can help the doctor or radiologist in the examine mammogram as a second opinion when perform a diagnosis. The system uses Support Vector Machine classification method, by perform classification of features value from digital mammogram. Where the features are generated from the segmentation and feature extraction.

This paper is limited to digital mammogram that used secondary data obtained from MIAS mammography database and limited to diagnoses abnormality of digital mammogram in the class of normal and abnormal.The paper aims: (1) find the region of interest or suspected area from mammogram image, (2) extract features from the mammogram image (3) classify normal and abnormal classes on mammogram image (4) determine the level of accuracy from classification results and (5) To find the influence of testing data on the level of accuracy from classification results

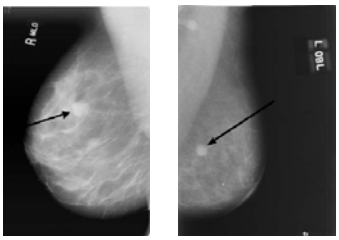
2. Literature Review

Breast cancer is a malignant neoplasm of the breast. A cancer cell has characteristics that differentiates it from normal tissue cells with respect to: the cell outline, shape, structure of nucleus and most importantly, its ability to metastasize and infiltrate. When this happens in the breast, it is commonly termed as 'Breast Cancer'. Cancer is confirmed after a biopsy (surgically extracting a tissue sample) and pathological evaluation or mammography exam [9]. Mammography is a specific type of imaging that uses a low-dose x-ray system to examine breasts. A mammography exam, called a mammogram, is used to aid in the early detection and diagnosis of breast diseases in women.

An x-ray (radiograph) is a noninvasive medical test that helps physicians diagnose and treat medical conditions. Imaging with x-rays involves exposing a part of the body to a small dose of ionizing radiation to produce pictures of the inside of the body. X-rays are the oldest and most frequently used form of medical imaging.

Two recent advances in mammography include digital mammography and computer-aided detection (1) Digital mammography, also called full-field digital mammography (FFDM), is a mammography system in which the x-ray film is replaced by solid-state detectors that convert x-rays into electrical signals. The electrical signals are used to produce images of the breast that can be seen on a computer screen or printed on special film similar to conventional mammograms. (2) Computer-aided detection (CAD) systems use a digitized mammographic image that can be obtained from either a conventional film mammogram or a digitally acquired mammogram. The computer software then searches for abnormal areas of density, mass, or calcification that may indicate the presence of cancer. The CAD system highlights these areas on the images, alerting the radiologist to the need for further analysis [2].

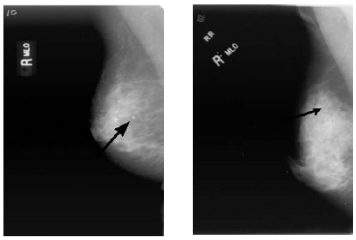
As we can see from the mammograms and diagrams, the breast tissue comprises of small intricate structures so pathology can easily be overlooked, especially if the film quality is not good. Not all structures are cancers. Many structures or macro (big) calcifications are often benign. Some basic forms of pathology and morphology presentations are : (A) Masses, differ from Densities because masses are seen on two views whereas densities are seen on one view only. Hence, two views of each breast to identify this abnormality. Masses with smooth rounded edge is generally a fluid-filled cyst that can be confirmed by an ultrasound and aspirated to relieve pain for the woman. A mass could be palpable (can feel it during a physical breast exam) depending on the size and proximity to the skin surface. Sometimes, it is very diffcult for radiologists to differentiate between a benign and a malignant mass off mammograms (as in Figure 1), so additional imaging modalities and/or biopsy may be required.



(a) (b)

Figure 1. Examples of (a) Benign Mass and (b) Malignant Mass

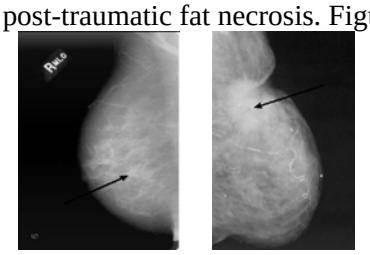
(B) Micro Calcifications, Calcifications are small calcium deposits that can be detected on a mammogram. Minute calcifications are called micro calcifications and bigger ones are called macro calcificiations. Artifacts on mammograms due to specs of dust may look like micro calcifications, but the difference is that these specs are bright and shiny whilst a micro calcification looks milky white.



(a) (b)

Figure 2. Examples of (a) Benign Calcification and (b) Malignant Calcification

(C) Spiculated lesions, This is by far the most definitive way to detect cancer. As a cancer cell proliferates, it shows up as a star-shaped or stellate lesion, with spiky lines radiating in all directions from a central region. A white star shape is characteristic of a malignant stellate lesion whereas the black star indicates a radial scar and post-traumatic fat necrosis. Figure 3.



(a) (b)

Figure 3: Examples of (a) Benign Spiculated and (b) Malignant Spiculated

The Mammographic Image Analysis Society (MIAS) is an organisation of UK research groups interested in the understanding of mammograms and has generated a database of digital mammograms. The database contains 322 digitised films and is available on 2.3GB 8 mm (ExaByte) tape. It also includes radiologist's "truth"-markings on the locations of any abnormalities that may be present. The database has been reduced to a 200 micron pixel edge and padded/clipped so that all the images are 1024x1024 [6]. In The MIAS database, mammogram image is divided into three classes,

that is glandular dense, fatty, and fatty glandular. In each class are subdivided into image of normal, benign and malignant. This is example of mammogram image from MIAS as shown in Figure 4.

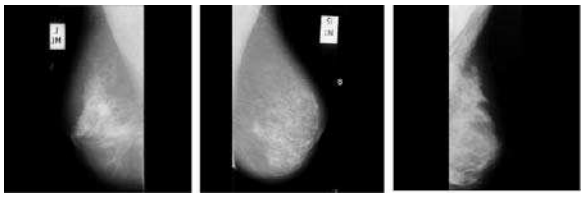


Figure 4. Mammogram Image From MIAS Dense Glandular (a) Benign (b) Malignant (c) Normal [6]

Graycomatrix creates the GLCM (Grey Level Co-occurrence Matriks) by calculating how often a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j [1]. Graycomatrix calculates the GLCM from a scaled version of the image. By default, if I is a binary image, graycomatrix scales the image to two gray-levels. If I is an intensity image, graycomatrix scales the image to eight gray-levels. We can specify the number of gray-levels graycomatrix uses to scale the image by using the 'NumLevels' parameter, and the way that graycomatrix scales the values using the 'GrayLimits' parameter.

The basic concept of Support Vector Machine (SVM) is actually a harmonious combination of computational theories that have been there before, such as margin hyperplane (Duda and Hart 1973, Cover 1965, Vapnik 1964, etc). A support vector machine constructs a hyperplane or set of hyperplanes in a high or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

3. Proposed Method and Discussion

The stages of breast cancer classification both normal and abnormal, which is divided into two processes, namely training process and testing process. Training process and testing process have the same stage. The first stage is image acquisition and then the second stage is do segmentation by finding the region of interest (suspected area of cancer). The third stage is extraction the features using GLCM. The fourth stage is perform mammogram image classification and the last is the result. These stage can be seen in Figure 5.

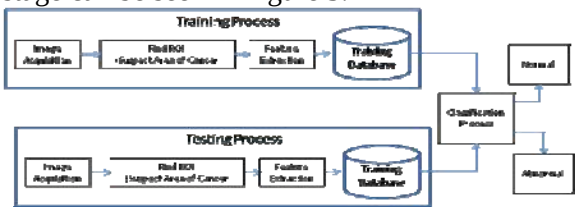


Figure 5. Architecture Design

3.1. Image Acquisition

The required data in this research are digital mammography image data that have normal case and abnormal case (benign breast cancer and malignant breast cancer). Based on these data requirements, the data used in this study is a standard data that are downloaded from the site abascus Mammography Database. In The MIAS database, mammogram image is divided into three classes, that is glandular dense, fatty, and fatty glandular. In each class are subdivided into image of normal, benign and malignant. Based on the data requirements that necessary to experiment, the sample data collected are as follows:

For training process, used 60 images. The image is consists of 30 normal images (10 dense glandular, 10 fatty, and 10 fatty glandular) and 30 abnormal images (15 benign and 15 malignant). Both of benign and malignant is taken 3 images from dense glandular, 3 images from fatty, and 4 images from fatty glandular.

For testing process, mammogram images performed in 3 cases. The first case is data used are taken inside from training data. The second case, data used are taken outside from training data and the third case, data used are taken inside and outside from training data. The data used for each case is 40 mammogram images. It consists of 20 normal images and 20 abnormal images (10 benign,10 malignant).

Image size used is 1024 x 1024 that means the image consists of 1024 columns and 1024 rows of pixels. So the total image consisted of 50873423 pixels, and the image has a PGM format that requires 8 bits per pixel with the

number of possible is 28 = 256 colors and possible color 0 (min) to 255 (max). The results can be seen in Figure 6. It is the original image from mammogram data (mdb028).

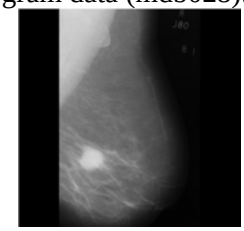
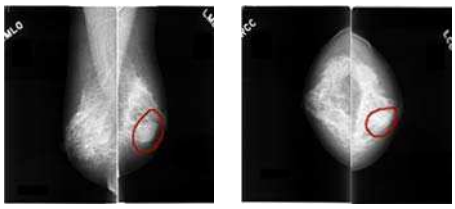


Figure 6. Original Mammogram Image (malignant mdb028.pgm)

3.2. Finding Region Of Interest (ROI)

A radiologist or doctor visually identify breast abnormalities by comparing mammography image taken from various points of view of image acquisition. Abnormalities were identified based on the characteristics of image, such as the left and right breast is not symmetrical. In addition, there is a similar shape with a circle that looks like a lump or there are areas where the image appears brighter that show the area has a higher intensity than the surrounding area as shown in Figure 7.



(a) (b)

Figure 7. Abnormal Mammography Image That Identified by Doctors

Visually, doctor recognize the left breast had abnormal characteristics. Therefore, doctor manually marked the areas that have abnormal characteristics as the suspected area of cancer (see Figure 7). In this research, The process of finding region of interest or suspected area is done by using morphology algorithms as shown in Figure 8. Morphology algorithm is used because it is able to separate abnormal area from normal area. Morphology operation used in this research is creation of structural elements operation, imtophat, imbothat, imsubtract, imcomplement, imextended, and imimposed.

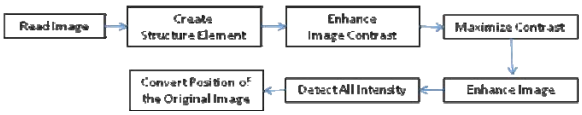


Figure 8. The Stages of Finding The ROI

3.3.2 .Create Structure Element (STREL)

The next step is create structure element, because it will apply dilation and erosion operations and it is an essential part of doing this is to create the structuring element that is used to probe the input image.

The structure element is a matrix consisting of only 0's and 1's that can have any arbitrary shape. Dilation and erosion functions is used to accept structure element objects and this objects is called STRELs. And strel function is used to create disk shape. Structural element is an operation that determines the shape and size of the addition and subtraction of pixels. Disk shape selected with consideration, every edge and curve of the object can be detected without none have passed. The size enlarged and shrinking expressed are determined by 10 pixels, because it can minimize the appearance of new information or missing information. Creation of structural elements are required for the next step segmentation using morphological operations.

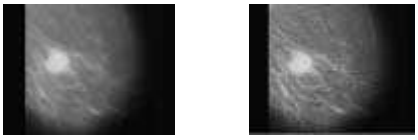
3.3.3. Enhance The Image Contrast

After structure of element is created, then performed the operation using imtophat and imbothat functions to enhance the image contrast. The operation was performed in order to displays detailed information from the invisible image.

3.3.4. Maximize The Contrast Between The Object and The Gaps

Next, to maximize the contrast between the objects and the gaps that separate them from each other by using subtract functions that added the top-hat image to the original image and then subtracts the bottom-hat image.

To maximize the contrast between the objects and the gaps that separate them from each other by using subtract functions that added the top-hat image to the original image and then subtracts the bottom-hat image. The result of maximize the contrast between the object and the gaps can be seen in Figure 9.



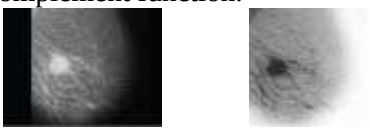
(a) (b)

Figure 9. (a) Original Image and (b) Maximize The Contrast Between The Object and The Gaps

3.3.5 Enhance The Image

The next step is convert object to photographic negative. It is particularly useful for enhancing white or gray detail embedded in a large, predominantly dark region.

The result of enhance the image can be seen in Figure 10. Enhance the image do with convert objects of Interest by watershed transform that detects intensity valleys in the image and also do enhanced the image by highlighting the intensity valleys using the imcomplement function. The negative of an image can be obtained with imcomplement function.



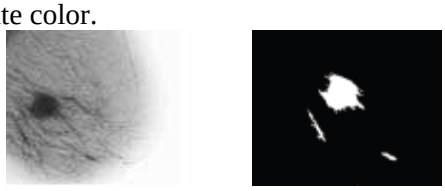
(a) (b)

Figure 10 (a) Maximize The Contrast and (b) Complement of Enhanced Image

3.3.6 Detects All The Intensity Valleys with Extended-minima Transform

After that, detects all the intensity valleys deeper than a particular threshold with the imextendedmin functions. This function used to mark areas that have a region minima and region maxima. The output of this function is a binary image.

This is the result (see Figure 11) of detects all the intensity valleys deeper than a particular threshold with the imextendedmin functions. The output of this function is a binary image. The result shows that region minima marked with black color and region maxima marked with white color.

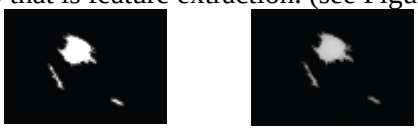


(a) (b)

Figure 11. (a) Complement of Enhanced Image and (b) Extended Minima Image

3.3.7 Convert The Position of The Original Image Intensity

Program above describe part of code to convert pixel values of Iemin (which has been done before) with binary values of 0 and 1. Conversion of binary value is conducted by changing the pixel value matrix 1 to be 0 and if the pixel is not 1 then converted into 1. Furthermore, the value is stacked into the original image. Where the position of the value 0 is converted into pixel value of original image and if the pixel is 1 then converted into 0. Therefore, the position of region of interest or segmentation found. This is the result segmentation mammogram image or Region of Interest (ROI). The segmentation results will be used for the next step that is feature extraction. (see Figure 12).



(a) (b)

Figure 12. (a) Extended Minima Image and (b) Region Of Interest

4.3 Feature Extraction

Feature extraction is very important part of pattern classification to identify texture in image, performed modeling texture as a two dimensional array gray level variation. This array called Gray Level Co-occurrence matrix. Five statistical measures such as contrast, correlation, energy, entropy, and homogeneity are computed based on GLCM. This research used direction 00 and distance 2. Feature extraction results below are training data by using 60 data (normal and abnormal). This data is a combination character of the background tissue, namely fatty, fatty glandular, and dense glandular.

The results of range between normal and abnormal mammogram image is obtained by finding the minimum and maximum values of all the data, then find range by maximum value minus minimum value and divide by the number range to be made. This research will be made 3 ranges, it because total data from all range can more easily be seen most clearly the difference. After getting the range value, then create the first range by adding the range value with the minimum value. Then, create the second range by

adding the first value with range value. The third value created by adding the second range value with range value. After that, from all the feature extraction data that have been obtained seen how the highest number of the first range, the second range and the third range. The highest number of each range used to see a significant difference between normal and abnormal from mammogram data.

Table 1. The Results Of Range Between Normal And Abnormal Mammogram Image

Features	Normal	Abnormal
Contrast	0.040565-0.238626	0.232185-0.439165
Correlation	0.896396-0.983610	0.928927-0.980383
Energy	0.868832-0.987895	0.745674-0.858751
Entropy	0.050317-0.304292	0.342372-0.593927
Homogeneity	0.989306-0.998193	0.980197-0.989456

SVM (Support Vector Machine) is a machine learning method that works on the principle of structural risk minimization in order to find the best hyperplane that separates two classes (normal and abnormal). The data used for this SVM is training data and testing data. In this research, testing data are divided into 3 groups. The first group, testing data were taken inside from training data. The second group, testing data were taken outside from training data. And the third group, testing data were taken inside and outside from training data. Testing data used are 40. Grouping is performed to see the accuracy from each group. The process of classification is performed to classify category of normal and abnormal from mammogram image.

From the classification results using testing data were taken inside from training data, percentage accuracy abnormalities value (normal or abnormal) is 85.00 %. From the classification results using testing data outside from training data, percentage accuracy abnormalities value (normal or abnormal) is 60.00%. From the classification results using testing data were taken inside and outside from training, percentage accuracy abnormalities value (normal or abnormal) is 67.50%. The accuracy of abnormalities (normal or abnormal) on mammogram data for each testing data have a different value. The first group have accuracy percentage of abnormalities by 85.00%. The second group have accuracy percentage of abnormalities by 60.00%. The third group have accuracy percentage of abnormalities by 67.5%.From the three groups, abnormalities accuracy values close to 100 percent is a group which using testing data were taken inside from training data by 85.00%.

4. Conclusion

In this research, diagnosis of Breast cancer in mammogram images using feature extraction and support vector machine is developed. The conclusion that the author retrieve from this research are: (1) Region of interest on mammogram image can be found using morphology algorithm. ROI or suspected area of cancer will be used to segmentation process. (2) Texture features have been proven to be useful in differentiating masses and normal breast tissues. The feature extraction method used is statistical method such as Grey level co-occurence matrix (GLCM). This research concentrate on statistical descriptors that include contrast, energy, homogeneity, entropy, and correlation of gray level values. (3) SVM (Support Vector Machine) is used to classify normal and abnormal classes on mammogram image. (4). The accuracy of abnormalities (normal or abnormal) in digital mammogram for each testing data have a different value. The first group have accuracy percentage of abnormalities by 85.00%. The second group have accuracy percentage of abnormalities by 60.00%. The third group have accuracy percentage of abnormalities by 67.5%.From the three groups, abnormalities accuracy values close to 100 percent is a group which is using testing data taken from training data by 85.00%.

This research is still necessary to development and improvement in the system. For the future, is expected to improve segmentation process (find the region of interest) by removing pectoral muscle and removing text noise from digital mammogram. Addition of GLCM features necessary to improve percentage of abnormality. Beside that, system can determine the level of severity (benign and malignant) from classification results. So it can help a doctor in diagnose breast cancer easily.

5. References

[1] Anonim, "Mammography", http://www.radiologyinfo.org/en/info.cfm?pg=mammo, June 2011.

[2] Anonim, "E Radiography", http://www.e-radiography.net/articles/mammo/mammo_introduction.htm, 2012

[3] A. Indrati, "Pengenalan Karakteristik Bentuk dan Batas Tepi Tumor Payudara," Ph.D. Dissertation, Gunadarma University, 2009.

[4] A. O. Malagelada, "Automatic Mass Segmentation in Mammographic Images," Ph.D. Dissertation, Universitat de Girona, April 2007.

[5] C.-I. I. Database, "Minimammographic Database", http://abacus.ee.cityu.edu.hk/imagedb/cgi-bin/ibrowser/ibrowser.cgi?folder=Medical_Image/mammogram/dense-glandular/, 2012.

[6] D. B. K. P. Rabi Narayan Panda and D. M. R. Patro, "Feature Extraction for Classification of Microcalcification and Mass Lession in Mammograms," 2009.

[7] D. S. Y. Ireaneus Anna Rejani, "Early Detection of Breast Cancer using SVM Classifier Technique," ser. 1, Pages 127-130, Publisher International Journal on Computer Science and Engineering, vol. 1(3), 2009.

[8] Dhika, "Apa Kabar Trend Kanker Payudara di Indonesia?", http://dhikatuy.blogdetik.com/2011/05/13/apa-kabar-trend-kanker-payudara-di-indonesia-2/, 2011.

[9] G, "Breast Cancer Incidence and Mortality Worldwide in 2008," http://globocan.iarc.fr/factsheets/cancers/breast.asp, 2008.

[10] M. H. Basim Alhadidi and H. N.Suleiman, "Mammogram Breast Cancer Image Detection using Image Processing Functions," ser. 2), Pages 217-221, Publisher Information Technology Journal, Vol. 6, 2007.

[11] P. P. D.Narain Ponraj, M.Evangelin Jenifer and J. Manoharan, "A Survey on The Preprocessing Techniques of Mammogram for The Detection of Breast Cancer," ser. 12, Pages 656-663, Month December, Publisher Journal of Emerging Trends in Computing and Information Sciences, Vol. 2, 2011.

[12] S. Basha and D. Prasad, "Automatic Detection of Breast Cancer Mass in Mammograms using Morphological Operators and Fuzzy C Means Clustering," pp. 704_709, 2009.

[13] S. B. Pradeep N, Girisha H. and K. K, "Feature Extraction of Mammograms," ser. 1, Pages 241-244, Publisher International Journal of Bioinformatics Research, Vol. 4, 2012.

Motion recognition of assistance for bed-wheelchair transferring from video

Masao Izumi† **Yurie Uno†** **Yukie Majima†**

†Graduate School of Engineering, Osaka Prefecture University, Japan

Abstract

Most of people who work as nurses have pain in the lower back. It is viewed that the mainly causes of pain in the lower back is the pose at assisting for bed-wheelchair transferring. In this paper, we propose a motion recognition system of assistance for bed-wheelchair transferring from video which made by home video to improve their assisting pose. For the purpose, we tried to extracting human region, fitting model, and extracting features of back and leg angle. After those steps, we investigated those features by comparing good assisting with bad assisting, and we gave a decision which indicates good or bad for input video. In the research, we confirm usefulness of our proposal method experimentally.

Keywords: motion recognition, bed-wheelchair transferring, Kinect, model fitting

1 Introduction

The nursing is classified into heavy handled transactions such as transfer of the patient, those with the risk of back pain onset often. Low back pain prevalence of nurses actually, differ according to actual survey, but most about 70%, there is a status quo that many nurses are suffering from low back pain. Further, so as going to face an aging society, it becomes a major problem that both of the helper-work nurses and patients would go on increasing. For a variety of reasons can be considered as a cause of back pain, the attitude of nursing operation is a major cause, especially the load on the lower back is applied to move assistance operation of bed-wheelchair transferring. In this study, we analyze the nursing operation, with the goal of building a learning support system to encourage the avoidance of behavior that cause low back pain. As a first step, we are trying to clarify the difference between the behaviors of the beginners and the experts. As a specific means, bed-wheelchair transferring nursing operation is performed the motion capture. We extract joint angles of nurses as a feature vector of the operation for the comparison on the time series variation.

As a method for extracting the angle of orientation, motion capture is typical. Various methods can be mentioned the capture technology, often measuring free operation is prevented in order to capture a combination of tracker to detect this with markers attached to the object, use range is limited. And there is the method which constructs a three-dimensional model from multiple cameras, and there is a technique using TOF (Time-of-flight) cameras, time-consuming setting equipment and is expensive. And it needs many difficulties for the user. In this study, we are trying to use Kinect[1], which is a less-expensive depth camera, for detecting three dimensional information. Following, we describe an overview of the process in Section 2, region detection method in Section 3, skeleton line detection method in Section 4, we describe how to obtain the feature values and the fitting method of the models in Section 5. In addition, it noted the Experimental results and discussion in Section 6, referred to a summary in Section 7 to the end.

2 Outline of Our Proposed Method

In this study, our system is going to obtain the joint angles of the nurse. For this sake, we are trying to detect human regions by using Kinect first. Detection of human regions uses a function which is one of functions implemented by OpenNI[2]. However, in the case of assistant operations, since the patient and the helper (= nurse) are detected without being distinguished, we cannot use this function of OpenNI directory in order to track nurse's motion. Once helper only region can be detected, we apply the skeleton line using a line detection operators Kasvand. Depth information from Kinect can be used for model fitting against detected skeletons, and joint angles of the body of the extracted nurse are extracted as feature vectors. Figure 1 shows the flow of the

overall processing.

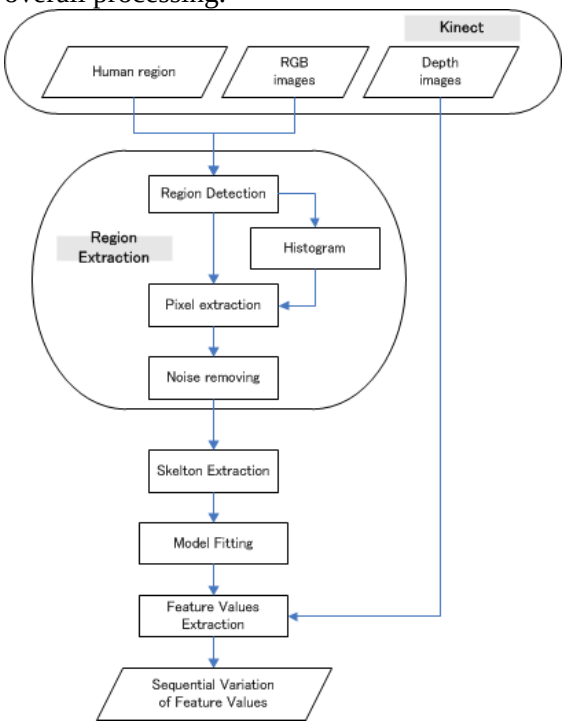
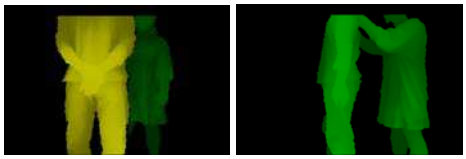


Figure 1 Flow of proposed method.

3 Region Detection of Helpers

In this section, we explain the method of region extraction processing of helpers. The OpenNI algorithms for recognizing a person region based literature such as [3] in advance is mounted. According to its manuals, it is possible to recognize plural people at the same time simultaneously, but this is a case where people with each other are not in contact. If two or more persons keep in touch to each other, system output the whole area which includes all persons as one area (Figure 2). In this study, a nurse and a patient tend to touch to each other, so Kinect output only one area which include both a nurse and a patient. In order to separate nurse area from joined area, the color of nurse's wear is different from that of patient's wear. (Figure 3).



(a)2 separate people (b) 2 joined people
Figure 2. Human Region Detection by Kinect.

(a) Input image (b) Extracted area
Figure 3. Human Region Detection by Kinect in case of Nursing works.

3.1. Region detection techniques

It is difficult to extract correct information from input images if there are some other areas which include non-target objects at the same time. In general, we are going to remove areas of non-target objects from whole extracted areas in order to handle the target object.
Background subtraction method is the simplest way to extract moving objects. In this method, input images are subtracting from the background image, and the subtracted value is higher than threshold, then its pixel becomes one of target areas. But this method can be applied to the case of simple background. If background is complicated, this method makes some noises (Figure 4).

(a) Background (b) input (c) subtraction
Figure 4. Examples of Background Subtraction.

There is also the method which continues to adapt to changes in modeling mixture distribution background, but the effect is only for a gradual change (Figure 5). And long-time non-moving object will disappear from foreground in this method (such as Figure 6).

(a) Input (b) subtraction
Figure 5. Examples of Background Subtraction in case of Shadow Effect.

(a) Input (b) subtraction

Figure 6. Examples of Background Modeling when a part of foreground was extracted as background.

So, we use the OpenNI region detection of the target first. And in the next, we apply the method for separate regions into the area in which only nurse exists.

3.2. Segmentation by using histograms.

Create a histogram of H (hue) component from HSV converted images. HSV converted images can be created by combining two images, one is the human regions which are extracted by Kinect, and another is RGB image which is taken through RGB camera of Kinect. It calculates the average from the values of the histogram, and stores the values were many occurrences than the average. *register* is the vector of 0 or 1, *value*(*i*) is the value of the histogram of the value *i*. (Equation 1)

$$\text{resister(i)} = \begin{cases} 1, & if \quad value(i) \geq ave \\ 0, & else \quad value(i) < ave \end{cases} \tag{1}$$

Figure 7 shows the result of nurse region extraction.

(a) Input (b) processed.

Figure 7. Examples of Region Partitioning.

Smooth area is not able to be extracted with this method, so we do correction processing by using expansion and contraction, labeling, and smoothing processing. Figure 8 shows the results of processing.

(a) Input (b) processed

Figure 8. Region Detection.

4 Skeleton detection

As a general method of skeleton detection, usually we use thinning process. But there is a problem that the processing time is greater, and some extra lines of skeleton tend to appear. So we use the operator Kasvand[4] for this purpose. However, since the topology is not guaranteed, so we performs expansion and contraction processes afterwards to avoid this problem.

4.1 Distance transformation

The pixels other than the background, it is referred to as *distance transformation* operations to replace the pixel value of the shortest distance from the background. Various distances are used in the distance transformation, but we show the distance transformation algorithm with four neighborhood distance here.

For input image $F = \{f(i, j)\}$, we raster scan to the lower right corner from the top left corner in the following rules in order to get the intermediate image $S = \{s(i, j)\}$. S is initialized to zero all the pixels.

$$s(i,j) = min\{ g(i,j), s(i-1,j)+1, s(i,j-1)+1\} \tag{2}$$

Where, $g\{(i,j)\}$ is

$$g(i,j) = \begin{cases} \infty, & if \quad f(i,j) = 1 \\ 0, & if \quad f(i,j) = 0 \end{cases} \tag{3}$$

Furthermore, the inverse raster scan to the upper left corner of the lower right corner based on the following rules: Now. Again, $H = \{h(i, j)\}$ is initialized to 0 all the pixels.

$$h(i,j) = min\{ s(i,j), h(i+1,j)+1, h(i,j+1)+1\} \tag{4}$$

Image $H = \{h(i,j)\}$ is the result of the distance.

4.2 Kasvand line detector

The line detection operator Kasvand is a line detection and operator with repeated emphasis of the local operation. To the S0, S45, S90, S135, respectively the value of the local product-sum operation for the four masks of 5 × 5 in Figure 9.

Where, coefficient □ is-1, coefficient ■ is 2, and coefficient ・is 0.

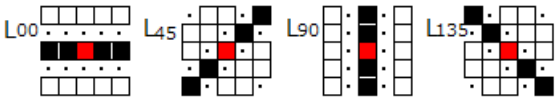


Figure 9. Templates of Kasvand Operator.

So we put S_{max} =Max(S^0, S^{45}, S^{90}, S^{135}), then output g is

$$g = \begin{cases} S_{max}, & if \quad S_{max} \geq 0 \\ 0, & else \end{cases} \tag{5}$$

Repeating above operation reaches the pixel which is larger than the certain threshold, then skeletons can be extracted. Figure 10 shows the resulting image.

(a) Input (b) processed

Figure 10. Skeletonized Images.

5 Model fitting and Feature value extraction

In this section, we apply the model fitting to the skeletons extracted in section 4, and extract feature values.

5.1 Model fitting

Model was used stick model consisting five sticks of the left and right thighs, the left and right lower legs, and waist.

First, the raster scan from the lower left under the following conditions for the skeleton lines extracted to determine the initial point.

$$start(x,y) = (i,j), \quad if \quad f(i,j) = 1 \tag{6}$$

Next, each stick is registered as a joint point when it reaches the registered length for each stick of the skeleton from the initial point determined above. Incidentally, the length of each stick is set manually, and as the tracking method, we use a technique of contour tracing this time. Five times of operations above reaches to the final model fitting. Figure 11 shows the tracking order model which was used this time.

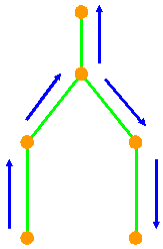


Figure 11. Fitting Model and Tracking orders.

Usually, contour extraction performs an operation such as Sobel filter, but in the case of a binary image, contours can be extracted so as following boundary lines of input image (contour tracing). The procedure is to determine the initial point first, and to explore the region pixels counterclockwise or clockwise from the target pixel. To extract the contour, we repeat this procedure until reaching to the initial point, and therefore it is generally used for fitting repeatedly specified number of times this time. Figure 12 shows the result image obtained by fitting the model to the actual image.

Figure 12. Results of Model Fitting.

5.2 Feature values extraction

Joint points obtained in 5.1 have the coordinates on the image, which are two-dimensional information. Therefore, using the depth information acquired at the time of shooting, converts the position of the real-world space position on the image to obtain a three-dimensional information. In terms of three-dimensional obtained information, we are trying to get joint angles as feature values. The acquisition method from three points adjacent to the center point is for calculating two three-dimensional vectors that are starting at the joint points. Where, calculated two vectors are $\vec{a} = (a_x, a_y, a_z)$ and $\vec{b} = (b_x, b_y, b_z)$

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z \tag{7}$$

$$|v| = \sqrt{v_x^2 + v_y^2 + v_z^2} \tag{8}$$

$$\cos\theta = \frac{\vec{a}\cdot\vec{b}}{|\vec{a}||\vec{b}|} \tag{9}$$

By using (7) to(9), it is possible to obtain the inner product of vectors *a* and *b* as the angle between vectors *a* and *b*.

Thus, the feature value 1 is the angle of left knee, the feature value 2 is the angle of the right knee, the feature value 3 is the angle of the left hip, and the feature values 4 is the angle of the right hip. Figure 13 shows these feature values.

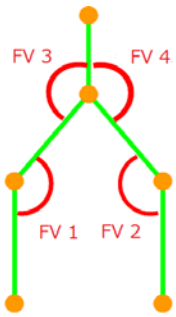


Fig. 13. Feature values.

6 Experiments and Discussions

6.1 Experimental conditions

As an experiment target images, we prepare two sets of expert samples in which expert nurse performed works for experiments, and four sets of beginner samples in which students of our university tried to perform works for experiments. Beginner samplers could see the expert's works just one time before performing their own works. In addition, we asked performers of experiments to dress in different colors wears in order to separate regions of nurses and patients, nurses wore the white dresses and patients wore the black dresses.

- target operation period: to hold in his arms and the patient that caused the upper body, until it down in a wheelchair (Figure 14) (in the nursing operation, this part is the heaviest load to his/her waist.)
- color data stream: 640 × 480 pixel 32 bit color
- depth data stream: 320 × 240 pixel unsigned 16 bit depth
- depth sensor width: 3.5 m to 1.2 m:λ
- frame rate: about 10 frame / second
- experimental image set:

  2 expert sample sets (33 sheets, 31 sheets)
  4 beginner sample sets (36 pieces, 31 pieces, 25 sheets, 20 sheets)

The experiments were carried out by creating a program based on the method of the video above, is shown in this paper. Incidentally, the language that was used to create the program is C + +.

Figure 14. Nursing works for experiments.

6.2 Experimental results

There are the result of charts which show chronological changes in the four feature values. Figure 15 shows the results of time series variation of the feature values of the experts. And Figure 16 shows the results of time series variation of features values of beginners. Vertical axes in both figures represent the degrees of angles of body joints, and horizontal axes represent the chronological change.

And areas of signed A, B and C in both figures show the areas which exist the certain differences between experts and beginners

.

Figure 15. Samples of experts' working postures.

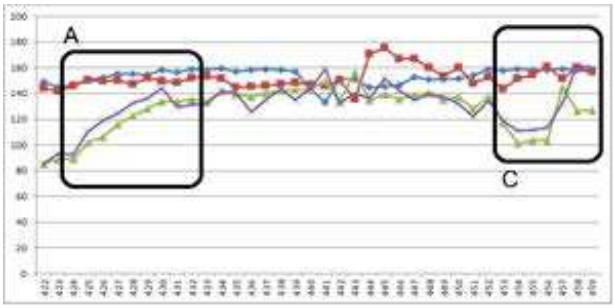


Figure 16. Samples of beginners' working postures.

6.3 Discussion

In this subsection, we describe three significant difference which are observed in the experiments. (1) In A part, all four values of experts become larger at the same time, but only two values (FV3 and FV4) of beginners become larger (FV1 and FV2 keep lower.) In A part, nurse should focus on his/her foot because he/she starts raising patients. Figure 17 shows the operation image of the A part of the actual.

(a) expert (b) beginner

Figure 17. Images of A-part.

(2) In B part, only expert samples show the characteristics. This part locates just after raising patients, and nurse should re-arrange the holding situation of patients. On the other hand, beginner samplers don't show the tendency of re-arrangement of holding patients in this part. Figure 18 shows the trends of these works.

Figure 18. Image of B-part (expert only).

(3) C part is a part of making the patient sit on the wheelchair. All four angles of experts go under 140 degrees. But those of beginners don't show the same patterns. It would show that experts try to put patients down by using all their bodies, but beginners tend not to use whole bodies for putting patient down into wheelchairs. Figure 19 shows the behaviors of experts in the C part.

(a) expert (b) beginner

Figure 19. Images of C-part.

In the future works, we will try to seek other kinds of tendencies of difference between experts and beginners.

7 Conclusions

In this paper, we have proposed the novel method of extracting nursing motion for bed-wheelchair transferring works by using Kinect. With experiments, we have shown the ability of our proposed method. In the future work, we will find more precise model fitting method and deep analysis of nursing motions. And we will gather much more test video sequences for farther investigation of the method.

References

[1] "Kinect-Xbox.com", http://www.xbox.com/ja-jp/Kinect, 2013/06/15 accessible

[2] "Introducing OpenNI", http://www.openni.org/, 2013/06/15 accessible

[3] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking", Proceedings of the IEEE Computer Science Conference on Computer Vision and Pattern Recognition (CVPR-99). 1999, Los Alamitos, IEEE, pp. 246–252.

[4] T.Kasvand. "Iterative edge detection". CGIP. 1979, vol. 4, pp. 279-286.

A Tennis Movie Retrieval System using Segmented-Swing Matching

Kaede Yamazaki† **Masato Kasahara‡** **Kosuke Takano†**

†Department of Information and Computer Sciences
Kanagawa Institute of Technology, Japan
‡Graduate School of Information and Computer Sciences
Kanagawa Institute of Technology, Japan
{ s1021159, s1385005}@cce.kanagawa-it.ac.jp, takano@ic.kanagawa-it.ac.jp

Abstract

In this paper, we present a prototype of a tennis movie retrieval system using a user's swing motion of a tennis racket. Our system leverages 3-axes acceleration data of swing motions obtained from a sensor device so that a user can retrieve the required tennis movies by intuitive swing motions with the sensor device. In our approach, the user's swing motion is segmented into three parts: pre-hit part, hit part, and post-hit part, for the swing-motion matching process. The experiment results show that by applying our segmentation method of swing motion to the search system, the retrieval accuracy can be improved with low computation cost.

Keywords: Tennis movie, Acceleration sensor, Information retrieval, Multimedia, Motion-based search

1 Introduction

With the advance of multimedia technology, a huge amount of movie contents available on PC are increasingly generated. Tennis movies are also widely provided from matches of professional players to lessons for beginners.

When a tennis learner searches for tennis movies for the purpose of improving his/her tennis play motions, he/she may input keywords that represent swing types of a tennis racket, which are some kinds of basic play motions in tennis. However, there are various swing motions according to stroke type such as forehand/backhand stroke, forehand/backhand volley, smash, and serve. Also, even in the basic strokes, there are several ones such as 'flat', 'drive', and 'slice'. Furthermore, in a tennis match, players use various anomalistic swing motions. Therefore, it is very hard for database creators to make a term-based index by using phrases such as 'strongly up-and-forward' and 'obliquely upward' that represent swing motions appeared in each tennis movie. In addition, it is difficult for searchers to use such phrases that represent the racket-swing motions in order to find specific tennis movies that they require.

Table 1. Example Descriptions of Tennis Strokes

Stroke Type	Description (from Wikipedia)
[Forehand Drive]	Topspin (Drive) is a property of a shot where the ball rotates as if rolling in the same direction as it is moving. Topspin on a shot imparts a downward force that causes the ball to drop, due to its interaction with the air. It can be generated by hitting the ball with an up-and-forward swing, with the racquet facing below the direction it is moving.
[Forehand Slice]	Slice is a shot such that the ball rotates backwards (as though rolling back towards the player) after it is hit. The trajectory of the shot involves an upward force that lifts the ball (see Magnus effect). While a normal hit bounces forward, when backspin shots bounce, they tend to bounce off the sides or even bounce up.

In order to address such a problem, our system leverages 3-axes acceleration data of swing motions obtained from a sensor device. The 3-axes acceleration data of swing motions are attached as metadata to the corresponding swing motions appeared in a tennis movie so that a searcher can retrieve the required tennis movie by intuitive swing motions with the sensor device, even in the case that it is difficult to make a term-based index

for each racket-swing motion.

In our approach, the user's swing motion is segmented into three parts: pre-hit part, hit part, and post-hit part, for the swing-motion matching process. By applying our segmentation method of swing motion to the search system, the retrieval accuracy can be improved with low computation cost.

In the implementation of this system, we use 3-axes acceleration sensor of Nintendo Wii system, which is a home video game machine. We confirm the feasibility of our system by experiments using our tennis movie search system.

2 Related Work

Motion tracking is an actively researched area [3, 4, 11] with the commercial success of the three major video game developers such as Microsoft, Nintendo, and Sony).

Sony PlayStation Move controller has a multi-color light source that, together with a Web cam, provides data of the controller's movement in the 3D space. Furthermore, accelerometer and inertial sensor within the handheld controller supply additional data to increase positional sensing accuracy [10].

Microsoft Kinect for Xbox 360 console is another Web cam based system; however, unlike Sony's Move controller, Kinect provides a device-less human machine interface using gesture and voice [8].

The Nintendo Wii remote controller, or the Wiimote, uses accelerometers and infrared detectors (together with the LED sensor bar) to sense the controller in the 3D space [9]. The add-on MotionPlus attachment is capable of supplying three axes of orientation change (or angular velocity) through its built-in gyroscopes.

Of the three, the Nintendo Wiimote is of particular interest as it presents a low-cost yet reasonably accurate means to capture motion or gesture data [1, 5]. Furthermore, the Wiimote can be set up easily and connected to a computer via a USB Bluetooth transceiver. However, the Wiimote has several limitations. The major one being the small "waggling" motions, intentional or not, could sometimes be interpreted as full movements. Though, this deficiency can be corrected by various means [12].

3 Tennis Movie Search System

3.1 System Overview

The proposed system searches a particular tennis video according to the movement of a swing as captured by the 3D acceleration sensor. Thus, users can find specific tennis videos by swinging their rackets imitating the swing motion of interest.

Our tennis instructional system retrieves a tennis video according to the input swing motion of the racket using the Wiimote 3D acceleration sensor in the following steps, as shown in Figure 1:

Step-1: Searchers do actually racket-swing-motion by exploiting Wii remote according to contents of required tennis movies.

Step-2: Motion-metadata extraction module extracts 3D acceleration data of the arm-motion in Step-1. The motion metadata extracted is used as query for tennis movie DB.

Step-3: Similarity computation module computes similarity of 3D acceleration data between query in Step-2 and each motion metadata linked with the corresponding tennis movie.

Step-4: Search IF show the ranking results based on the similarity values in Step-3.

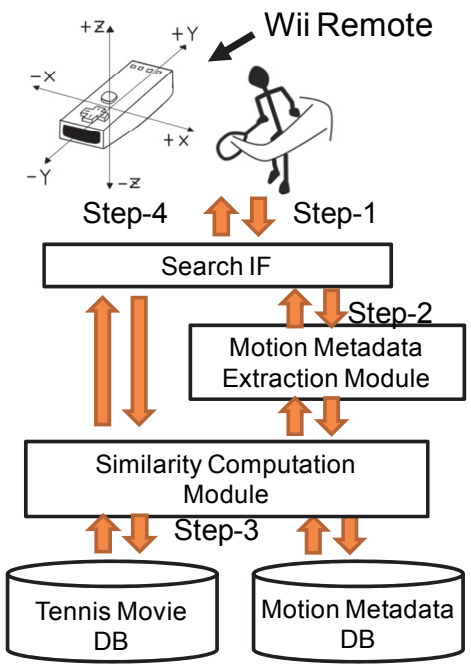


Figure 1. Retrieval of Tennis Video

3.2 Basic Similarity Computation

Firstly, we define motion metadata that represent 3-axes (x, y, z direction) acceleration data at each time be P, Q as follows:

$$\begin{aligned} P &: p_1(a_{x1}, a_{y1}, a_{z1}), p_2(a_{x2}, a_{y2}, a_{z2}), \\ &\qquad \cdots, p_n(a_{xn}, a_{yn}, a_{zn}) \\ Q &: q_1(b_{x1}, b_{y1}, b_{z1}), q_2(b_{x2}, b_{y2}, b_{z2}), \\ &\qquad \cdots q_n(b_{xm}, b_{ym}, b_{zm}) \end{aligned} \tag{1}$$

For the similarity calculation, we leverage two measures using Euclidian distance and DTW-based distance.

[Basic distance]
As a basic distance, we apply the simplest similarity measure using Euclidian distance. The similarity score D_{euclid} between P and Q is calculated as summation of Euclidian distance as follows:

$$\begin{aligned} D_{euclid}(P,Q) &= p_1 q_1 + p_2 q_2 + \cdots + p_k q_k \\ &= \sum_{p=1}^{k} \sqrt{(a_{xp} - b_{xp})^2 + (a_{yp} - b_{yp})^2 + (a_{yp} - b_{yp})^2} \\ k &= \begin{cases} m & (m \le n) \\ n & (m > n) \end{cases} \end{aligned} \tag{2}$$

[DTW-Based Distance]
In another choice of the similarity measure, we apply a distance based on a DTW (Dynamic Time Warping) algorithm [13].

$$D_{dtw}(P,Q) = \mathrm{dtw}(P,Q) \tag{3}$$

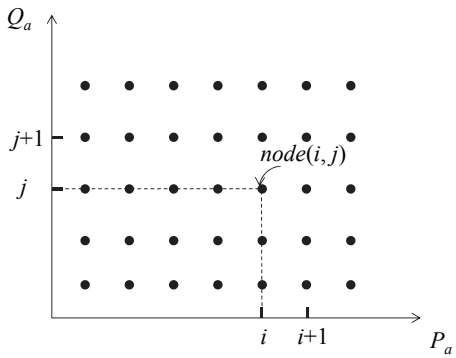


Figure 2. Nodes for applying DTW algorithm

For applying the DTW algorithm, we define a node (i, j) as shown in Figure 2 and the node cost of each node (i, j) in Eq. (4). Using the node cost in Eq. (4), the basic DTW algorithm employed is shown in Figure 3.

$$\begin{aligned} \mathrm{node_cost}(i,j) &= p_i q_j \\ &= \sqrt{(a_{xi} - b_{xi})^2 + (a_{yi} - b_{yi})^2 + (a_{zi} - b_{zi})^2} \end{aligned} \tag{4}$$

```
%Initialization
D(1,1) = node_cost(1,1);

for i = 2:I
D(i,1) = D(i-1,1)+node_cost(i,1);
end

for j = 2:J
D(1,j) = D(1,j-1)+node_cost(1,j);
end

%Main Loop
for i = 2:I
for j = 2:J
D(i,j) = min(D(i-1,j-1) + node_cost(i,j),
D(i-1,j) + node_cost(i,j),
D(i,j-1) + node_cost(i,j));
end %for j
end %for i
matching_cost = D(I,J);
```

Figure 3. Basic DTW algorithm

4. Segmented Swing Matching

4.1 Swing Segmentation

We propose a segmentation method of swing motion data in order to realize more flexible swing-motion matching. In this method, swing motion data is partitioned into three parts: (1) pre-hit part, (2) hit part, and (3) post-hit part as shown Figure 4.

There are several types of tennis swings. For example, when we distinguish forehand drive and forehand flat swings, the analysis part to be focused is a 'hit' part, because in the forehand drive shot, a strong upper forward spin is given at the moment (Figure 5 (a)). Also, in swing of volley, the length of post-hit part is short, on the contrary, the length of post-hit part is long in swing of serve. Thus, according to the swing type, the analysis part to be focused would be different for identifying the swing motion (Figure 5 (b)).

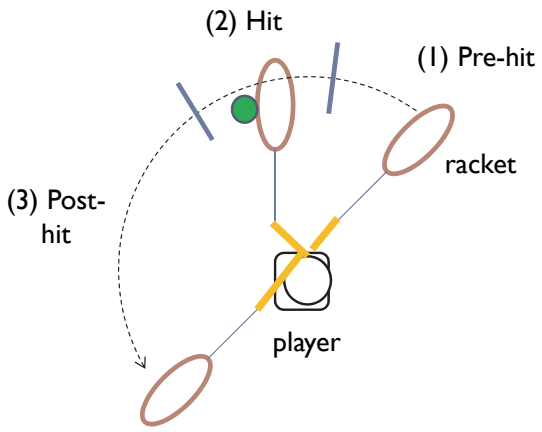


Figure 4. Segmentation of swing motion

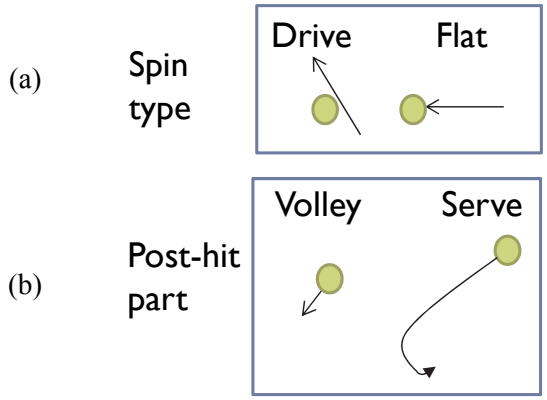


Figure 5. Example of character of swing motion: (a) spin type and (b) swing part

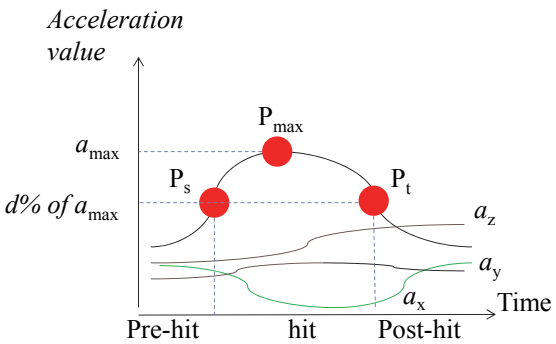


Figure 6. Segmentation points

Our segmentation method of swing motion is executed in the following steps:

Step-1: Find a point p_{max} with a max acceleration value a_{max} for swing motion P: $p_1, p_2, \ldots, p_n$ (Figure 6)

Step-2: Find points p_s and p_t, in which acceleration values are d % of a_{max}, forward and backward from p_{max}, respectively.

Step-3: Devide sequence of points into three segments P_{pre}, P_{hit}, P_{post} as follows:

$$P_{pre}: p_1, p_2, \ldots, p_s$$
$$P_{hit}: p_{s+1}, p_{s+2}, \ldots, p_t$$
$$P_{post}: p_{t+1}, p_{t+2}, \ldots, p_n$$

4.2 Swing Matching

In applying both basic distance and DTW-based distance, the similarity score is calculated as the summation of distances from each segmented part. In the following equations (5) and (6), w_{pre}, w_{hit}, and w_{post} are weight values assigned to each segmented part.

[Basic distance]

$$\begin{aligned} \text{SimiralityScore}_{basic} = & \; w_{pre} \times D_{euclid}(P_{pre}, Q_{pre}) \\ & + w_{hit} \times D_{euclid}(P_{hit}, Q_{hit}) \\ & + w_{post} \times D_{euclid}(P_{post}, Q_{post}) \end{aligned} \quad (5)$$

[DTW-based distance]

$$\begin{aligned} \text{SimiralityScore}_{basic} = & \; w_{pre} \times D_{DTW}(P_{pre}, Q_{pre}) \\ & + w_{hit} \times D_{DTW}(P_{hit}, Q_{hit}) \\ & + w_{post} \times D_{DTW}(P_{post}, Q_{post}) \end{aligned} \quad (6)$$

5 Experiments

In the experiment, to illustrate the feasibility of a tennis movie search system, preliminary testing and validation have been carried out.

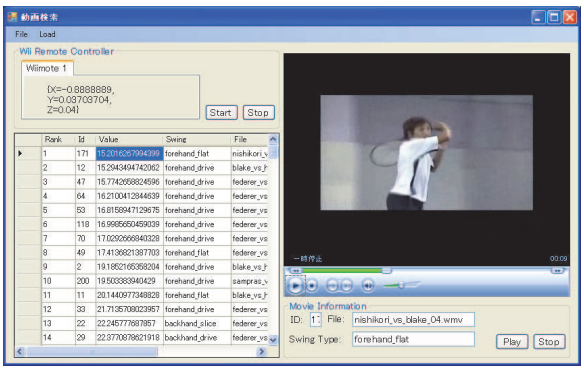


Figure 7. Snapshot of Our System

5.1 Environment

Figure 7 shows the snapshots of our prototype system. Forecasting a tennis movie retrieval system, middleware and user-interface can be developed using C# and installed on the widely available PCs. The Wiimote can be set up easily and connected to a computer via a Bluetooth

transceiver. The experimental platform for detecting tennis moves includes a laptop and a Wiimote. The motion sensor data from Wiimote are directly transferred to our system and processed for the motion data matching.

In the experiment, we used 201 tennis movies. An experienced tennis player manually configures motion metadata for each tennis movie in advance. One motion of metadata is attached for one tennis movie. In addition, we classified tennis movies into 57 patterns of swing motions (19 swing types x 3 speeds (slow, middle, and quick)) as shown in Table 2.

Table 2. Number of Tennis Movies in Each Classification
F: Forehand, B: Backhand
S: Slow, M: Middle, Q: Quick

	Swing Type	S	M	Q	Total
1	F drive	1	28	15	44
2	F slice	0	1	0	1
3	F flat	0	4	20	24
4	F drop	3	0	0	3
5	B drive	2	26	6	34
6	B slice	1	5	0	6
7	B flat	1	4	6	11
8	B drop	3	0	0	3
9	F low volley	3	0	0	3
10	F middle volley	9	0	0	9
11	F high volley	2	1	0	3
12	F drive volley	0	0	5	5
13	B low volley	3	0	0	3
14	B middle volley	7	1	0	8
15	B high volley	3	1	0	4
16	Smash	0	3	6	9
17	Flat serve	0	1	21	22
18	Slice serve	0	6	0	6
19	Spin serve	0	3	0	3
	Total	38	84	79	201

5.2 Evaluation Method

We examine precision-recall rate for retrieval results using basic distance and DTW-based distance, respectively, as follows:

Precision rates
= (# of the correct tennis movies in the top-10) / 10

Recall rates
= (# of the correct tennis movies in the top-10) / (# of the correct tennis movies)

In addition, we examine precision-recall rate in both case of (a) considering about swing speed and (b) not considering about it. Computation time for calculating the similarity using basic distance and DTW-based distance is also discussed.

5.3 Result

Figure 8 shows the precision-recall graph of retrieval results using basic distance and DTW-based distance. The dot line graphs mean the case of considering swing speed for recall-precision calculation. Table 3 shows the result of computation time. From these results, we can confirm that accuracy of DTW-based distance is better, but, it took much computation time.

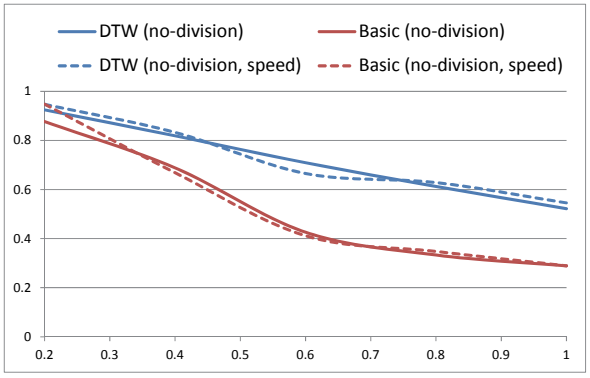


Figure 8. Recall-precision graph: basic distance and DTW-based distance

Table 3. Computation time: basic distance and DTW-based distance

	AverageTime (s)
Basic (no-division)	0.0572
DTW (no-division)	41.3871

Figures 9 and 10 show the recall-precision graphs for basic distance. Figure 9 shows the case of not considering the swing speed, and Figure 10 shows the case of considering the swing speed.

We can confirm that retrieval accuracy changes according to segmentation threshold. In this result when thold = 50%, we can get best performance result.

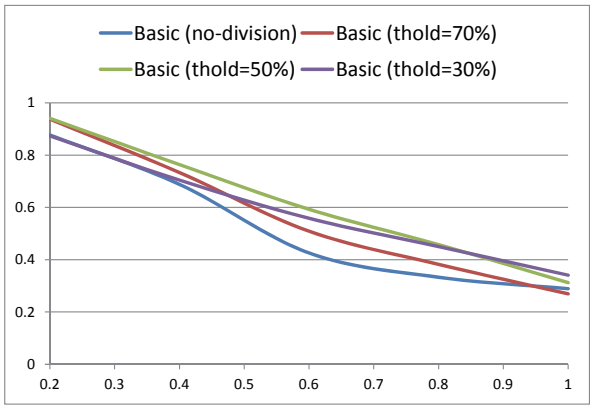


Figure 9. Recall-precision graph: basic distance (no-consideration of swing speed)

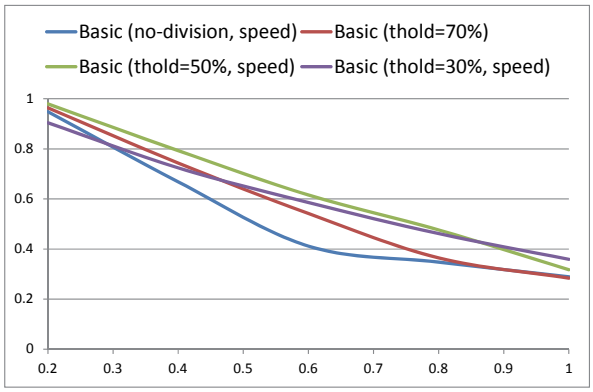


Figure 10. Recall-precision graph: basic distance (consideration of swing speed)

Figures 11 and 12 show the recall-precision graphs for DTW-based distance. The detail numerical values of each Figure are given in Tables 4 and 5. Figure 11 and 12 shows the case of not considering the swing speed and the case of considering the swing speed, respectively. Unlike our expectation, we can find little change by segmentation of swing data. This is because due to time stretching capability in the matching process, DTW don't need to detect boundary of three swing segments: pre-hit part, hit part, post-hit part.

However, the result of Table 6 shows that computation time is reduced by segmentation of swing data. This is because computation time by DTW highly depends on the length of time-series data, and by segmentation, length of each segmented swing data is shorten so that computation time by DTW can become lower.

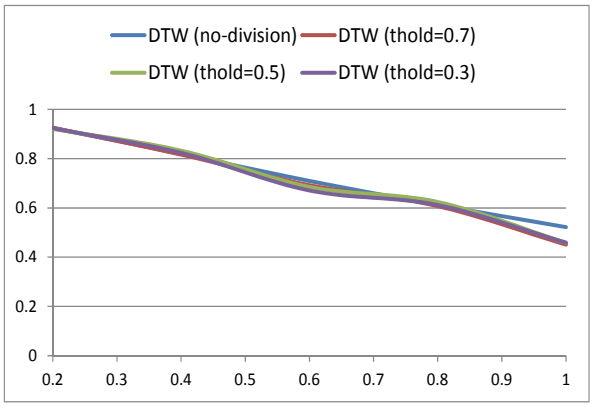


Figure 11. Recall-precision graph: DTW-based distance (no-consideration of swing speed)

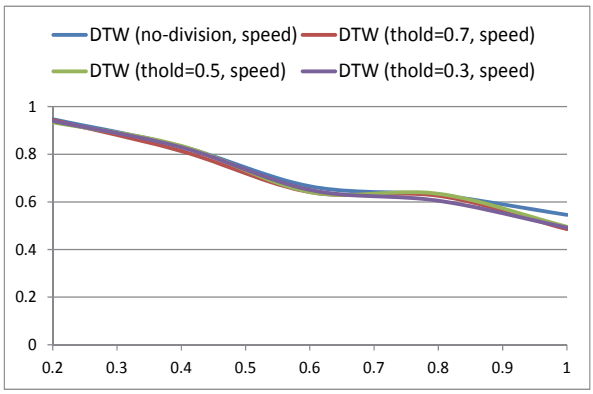


Figure 12. Recall-precision graph: DTW-based distance (consideration of swing speed)

Table 4. Numeric values of Figure 11

Recall	Precision			
	DTW (no-div)	DTW (th=0.7)	DTW (th=0.5)	DTW (th=0.3)
0.2	0.925	0.925	0.920	0.923
0.4	0.819	0.816	0.833	0.823
0.6	0.709	0.692	0.685	0.671
0.8	0.613	0.607	0.624	0.612
1.0	0.522	0.451	0.458	0.459

Table 5. Numeric values of Figure 12

Recall	Precision			
	DTW (no-div, speed)	DTW (th=0.7, speed)	DTW (th=0.5, speed)	DTW (th=0.3, speed)
0.2	0.947	0.943	0.934	0.939
0.4	0.833	0.813	0.835	0.829
0.6	0.665	0.641	0.641	0.650
0.8	0.628	0.626	0.634	0.605
1.0	0.546	0.486	0.496	0.492

Table 6. Computation time: DTW-based distance

	Average Time (s)	Average # of tuples
DTW (thold=50%)	18.7981	Pre:105, Hit:42, Post: 23
DTW (no-division)	41.3871	167

Figures 13 and 14 are shown for comparison of accuracy of the search results. From these results, we can confirm that our segmentation method of swing data improves the search accuracy. The accuracy of basic distance is less than DTW-based distance, however: the computation time of basic distance is much faster as shown in Table 3.

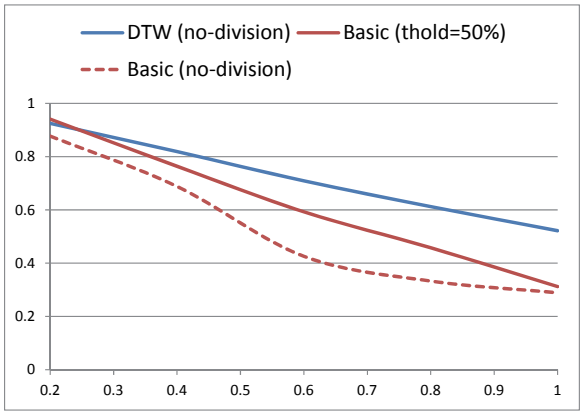


Figure 13. Comparison of accuracy (no-consideration of swing speed)

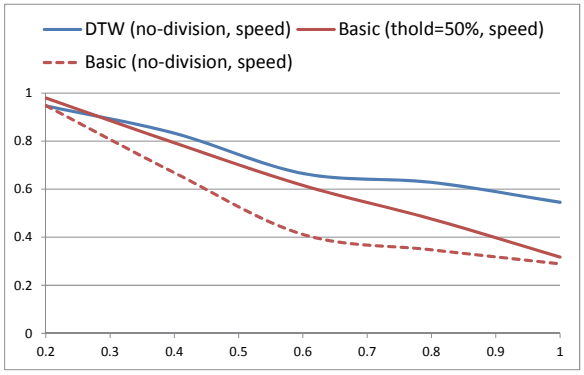


Figure 14. Comparison of accuracy (consideration of swing speed)

6 Conclusion

In this paper, we have presented swing-motion segmentation technique for a tennis movie retrieval system. In the implementation, we use 3-axes acceleration sensor of Nintendo Wii controller to capture the user's swing motions.

We conducted several experiments by applying the basic distance similarity and the DTW-based distance similarity for the matching and identification of tennis swing data. In the experiment, our segmentation method of swing data shows enough results to improve retrieval accuracy with low computation cost in the result using basic distance. In addition, we confirmed that our method reduced the computation time for the swing matching using DTW-based distance. These results show the feasibility of our system to retrieve the tennis movies as an initial step to realize a practical system.

Despite of Wiimote incapability to capture correct sensor values for quick motions, our system could retrieve the correct tennis movies in some queries. In addition, we used only 3-axes acceleration values, however; by considering the direction (yaw, pitch, roll) of the racket surface, we think that the search precision will be improved better.

In our future work, we will realize a learning function for adopting individual's character of the swing motions in our search system.

Acknowledgment

We appreciate Dr. Kin Fun Li at University of Victoria, Canada for his much useful advice to this study.

References

[1] Bradshaw, D., and K. Ng. *Tracking Conductors Hand Movements Using Multiple Wiimotes*, International Conference on Automated Solutions for Cross Media Content and Multi-Cannel Distribution, pp. 93-99, 2008.

[2] Diehl, G.M., J.f. Lynch, and D. Bastone. *Sports Instruction system and Method*, available at http://www.faqs.org/patents/app/20080293023, October 20, 2010.

[3] Hay, S., J. Newman, and R. Harle. *Optical Tracking Using Commodity Hardware*, IEEE International Sympoisum on Mixed and Augmented Reality, Cambridge, UK, pp. 159-160, 2008.

[4] Hoffman, M., P. Varcholik, and J.J. LaViola Jr. *Breaking the Status Quo: Improving 3D Gesture Recognition with Spatially Convenient Input De-*

vices, IEEE Virtual Reality, pp. 59-66, 2010.
[5] Lee, J.C. *Hacking the Nintendo Wii Remote*, IEEE Pervasive Computing, Volume 7, Issue 3, pp. 39-45, 2008.
[6] Li, K.F., and M.G. Johnson. *Capturing Motion Data with the Wiimote*, NEWS Technical Report, Department of Electrical and Computer Engineering, University of Victoria, Canada, August 2010.
[7] Lo, C.-Y., H.-I. Chang, and Y.-T. Chang. *Research on Recreational Sports Instruction Using an Expert System*, LNCS 5820, J. Liu et al. editors, pp. 250-262, 2009.
[8] Microsoft XBOX, available at http://www.xbox.com, October 20, 2010.
[9] Nintendo Wii System, available at http://wii.com, October 20, 2010.
[10] Sony PlayStation, available at http://www.playstation.com, October 20, 2010.
[11] Wang, Y., et al. *Using Human Body Gestures as Inputs for Gaming via Depth Analysis*, IEEE International Conferene on Multimedia and Expo, pp. 993-996, 2008.
[12] Wingrave, C.A., et al. *The Wiimote and Beyond: Spatially Convient Devices for 3D User Interfaces*, IEEE Computer Graphics and Applications, pp. 71-85, Volume 30, Issue 2, 2010.
[13] Sakoe, H., and S. Chiba. *Dynamic Programming Algorithm Optimization for Spoken Word Recognition*, IEEE Transactions on Acoustics, Speech and Signal Processing, Volume 26, Issue 1, pp. 43-49, 1978.

A Study on the Development of an Occupancy Simulator for Elderly Monitoring

Tomoharu Nakashima[†] **Yasuko Maekawa**[‡]
Youhei Kuroda[†] **Ayaka Dobashi**[†] **Yumi Higuchi**[*]
[†]Graduate School of Engineering, Osaka Prefecture University
[‡]College of Sustainable System Science, Osaka Prefecture University
[*]Graduate School of Comprehensive Rehabilitation, Osaka Prefecture University
tomoharu.nakashima@kis.osakafu-u.ac.jp
yasuko_m@kis.osakafu-u.ac.jp
{youhei.kuroda, ayaka.dobashi}@cs.osakafu-u.ac.jp
yumi@rehab.osakafu-u.ac.jp

Abstract

In this paper, a simulator of an elderly occupant is developed. The developed simulator generates the pattern of transition between multiple areas inside the house. The generation of the transition pattern is based on a typical transition pattern that is identified from a set of observed data in a real experimental environment. The validity of the generated occupancy is investigated using some matching measures.

Keywords: Elderly care, Simulation, Occupancy profile, Sensory data

1 Introduction

With the increase of the elderly people especially in Japan, the number of elderly people with a physical disability and dementia increases, too. Considering this social situation, the support by family and trained caregivers is necessary for the elderly person with disorders so that the elderly people can receive support from them for their daily life and protecting from their physical accidents. However, in the current Japanese society, it is difficult to secure enough number of trained people that can accommodate rapidly increasing elderly people. Therefore, the expectation for the automatically monitoring systems has been raised so that the elderly people with dementia, a physical disability become independent as safely as possible.

The automatic monitoring systems can also reduce the burden of their families and caregivers, and therefore can help settle the issue of the increasing demand for elderly care. Smart homes that are equipped with sensors for monitoring the state of the occupants is one of active research areas. The sensory data that are collected through infra-red motion sensors, pressure sensors, door sensors, etc. are integrated and used for the analysis of the occupants. It is desirable to monitor residents while protecting their privacy in their rooms using non-invasive sensors that do not give them unnecessary stress. In addition, the automatic monitoring system can call the family or caregivers when any abnormality happens to the residents.

It is necessary to model the normal life pattern of the occupants in order to automatically perform the abnormality detection. Once the normal life pattern is successfully obtained, the system checks if the occupants are normal or not by checking the difference between the current life pattern and the normal one. Modelling the normal life pattern requires a number of occupancy data that are obtained through the equipped sensors. However, it is difficult to obtain enough amount of the observed data as it takes long time to collect sensory data from one particular occupant.

In this paper, an occupancy simulator in the smart home environment is developed. The simulator virtually plays a role of the occupant in a house and help generate the sensory data. The occupancy simulator can generate the simulated data much faster than real observed data as it can accelerate the simulated time in the virtual smart home environment.

First, the procedure for generating occupancy data by the simulator is explained. The occupancy simulator in this paper needs a typical occupancy pattern called an occupancy profile that is used as a base for generating simulated occupancy data. Then, the validity of the simulated

occupancy data is examined by comparing them with the observed data that are obtained from the real smart house environment. As a validity measure of the simulated occupancy data, a concordance rate between the observed data and simulated data is used. In order to enhance the quality of the simulated occupancy data, the occupancy simulator is extended so that multiple occupancy profiles are allowed that represent the observed real data as a group of the typical occupancy patterns.

2 Occupancy simulator

The occupancy simulator is aimed at reproducing the transition of the observed occupant in a real house. In the simulator, a virtual occupant moves around the rooms for 24 hours (i.e., one day). The virtual occupant moves five places (a bedroom, a bathroom, a living, a kitchen and outdoors). The duration time of each place is decided using a normal random number. The occupancy simulator generates the occupancy data based on an occupancy profile. The occupancy profile is assumed to be obtained before its execution. The occupancy profile includes not only a series of transition between the places but also the duration of being in each place [1]. The following is the procedure of the occupancy simulator:

1. Set the initial place of the day as the first place in the occupancy profile.

2. Determine whether the virtual occupant follows the occupancy profile for the next place or not with some probability.

3. If the next place was determined to follow the occupancy profile in **2**, the duration in the place is determined using a normal distribution with the average in the profile and the variance of 10% of the average. The random number is generated using this normal distribution and rounded off the decimal first place to obtain an integer duration time. Go to 5

4. If the next place was determined not to follow the occupancy profile in **2**, the next place is determined according to the transition matrix. In this case, the duration of the next room is fixed to ten minutes.

5. **1 – 4** are continued untill a grand total of the duration time reaches on (1st) for 1,440 minutes or it is at the end of a given typical occupancy pattern.

If all transitions in the occupancy profile is used, the virtual occupant goes back to the the first room until the 1,440th minute of the day (i.e., the end of the day). The typical occupancy profile is less likely to be completely used until its final transition if the virtual occupant does not follow the sequence in the step **enu:determine** many times. This is because when the virtual occupant does not follow the occuant profile, an extra 10 minutes is incurred, which prolong the transition pattern. Therefore, we consider three methods for getting back to the occupancy profile when the virtual occupant does not follow the occupancy profile. The followings are the detail of the three methods:

Method 1 This method is the original and simplest way. The virtual occupant goes back to the interrupted transition in the occupancy profile.

Method 2 This method subtracts ten minutes (i.e., the duration in the current place) from the duration at the next place in the occupancy profile. If the duration of the next place becomes zero as a result of the subtraction, the next place after the next is used when the virtual occupant get back to the place transition according to its occupancy profile.

Method 3 When the action not following the typical occupancy profile changes into a action following it, if the time is before the expected start time of the action following the given typical occupancy profile, it does the process of method 1. If the time is after that, it does the process of method 2.

3 Acquisition of the typical occupancy profile from observed data

3.1 Observed data

Acquisition of the typical occupancy profile is carried out from observed data. There are the observed data for eight days, and they are collected by the sensor put in the occupant's house. The sensors detect the movements of any objects in its covered area, and report the sensor ID the

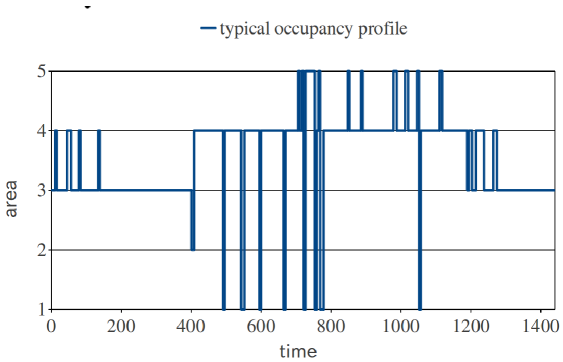


Figure 1. The occupancy profile obtained from the observed data for eight days.

time when it captured the object. In the preprocessing of the observed data, if the reaction from a particular sensor lasts for some seconds, it is considered as bungle and is excluded from the final observed data. The obtained sensor data were taken in a house in U.K. The sensors are installed at the front door, the back door, a bedroom, a bathroom, a lounge, and a kitchen. If a door sensor (at either the front door and the back door) is being reacted, the occupant is considered to go out of the house.

3.2 Acquisition of Occupancy Profiles

This subsection explains how an occupancy profile is obtained from observed data for eight days. Each of the eight observed data are scanned with a resolution of one minute. For each minute, the majority place among the eight days is selected as the typical place for that particular minute. This procedure is conducted for for 1,440 minutes (i.e., for a day). The duration time for each place in the occupancy profile is determined as the average time over the eight days. An occupancy profile acquired by the above procedure is shown in Fig. 1.

Next, we calculated the concordance rate between the obtained occupancy profile and the observed data. Table 1 shows the momentary concordance rates of the typical occupancy profile and observed data. Table 1 shows the concordance rates of the occupancy profile and the observed data for each minute. The concordance rate is low around 30–55% during the day. It is possible that eight days are not enough for obtaining an occupancy profile with a good quality and high concordance rates.

Table 1. Concordance rates between the occupancy profile and the observed data.

Day	Concordance rate
1	77.6
2	69.9
3	68.8
4	70.0
5	79.4
6	68.2
7	60.1
8	68.7

Table 2. Concordance rate between the occupancy profile and the observed data.

Day	Concordance rate
1	39.6
2	47.9
3	50.0
4	33.3
5	54.2
6	52.1
7	29.2
8	29.2

3.3 Acquisition of a Transition Matrix

The transition matrix is used when the virtual occupant does not follow the occupancy profile. The matrix shows the probability of each transition from each place. In a similar way to the case in the occupancy profile, the transition matrix is also obtained from the observed data. For all the observed data for the eight days, the transition is recorded and counted. Let us denote c_i^j as the number of the transition from place i to place j. The element a_i^j of the transition matrix that shows the probability of the transition from place i to place j can be calculated as follows:

$$a_i^j = \frac{c_i^j}{\sum_j c_i^j}, \tag{1}$$

Table 3 shows the transition matrix that was obtained from the observed data. One can see that the probability of the transition to the lounge and the kitchen is higher than the other places. Thus, the occupant tended to move the lounge and kitchen with a high probability.

Table 3. Transition matrix that is obtained from the observed data.

area	outside	bedroom	bathroom	lounge	kitchen
outside	-	0.02	0.06	0.35	0.57
bedroom	0.04	-	0.37	0.23	0.36
bathroom	0.08	0.31	-	0.42	0.19
lounge	0.14	0.06	0.14	-	0.66
kitchen	0.23	0.11	0.07	0.59	-

4 Numerical Experiments

4.1 Calculating Concordance Rates

In order to check whether the developed simulator successfully generate appropriate occupancy patterns or not. Numerical experiments are conducted to investigate the quality of the generated occupancy patterns. The concordance rate is used as a performance measure of the simulator. The concordance rate is the proportion of the transition or the occupancy that matches with the real observed data. Two types of concordance rates are considered to investigate the quality of the generated occupancy patterns in detail. One is temporal concordance rate and the other is transitional concordance rate. The temporal concordance rate shows how many minutes out of 1,440 (i.e., a day) the occupied places by the virtual occupant agree with that of the real observed data. On the other hand, the transitional concordance rate checks how many transitions (that is, a pair of places that the occupant moves from and to) agree with the observed data. If the concordance rates of the occupancy profile and the generated occupancy patterns by the simulator, it can be said that the generated occupancy patterns are appropriate. The following is the description of the two concordance rate:

Temporal concordance rate

If a place of typical occupancy profile within a certain time interval also appears in the generated/observed data, it is called temporal concordance.

Transitional concordance rate

If the transition between the occupancy profile within a certain time interval also appears in s seen in the generated/observed data, it is called transitional concordance.

4.2 Occupancy Simulation

The occupancy simulation is conducted using the obtained occupancy profile in Subsection 3.2 and the transition probability matrix in Subsection 3.3. The uncertainty factor that specifies the probability that the virtual occupant does not follow the occupancy profile is set to from 0% to 50% with a step of 5%. For each uncertainty factor, Method 1, 2, and 3 are also applied. In addition, the occupancy pattern is generated without the uncertainty factor but with the transition matrix that in this experimental setting the virtual occupant totally follow the transition matrix any time.

Next, Methods 1 – 3 are applied for each experimental settings and 100 occupancy patterns are generated by the developed simulator. For the generated occupancy patterns, the concordance rates are calculated. The table of average of 100 times of the concordance rate of the simulation data generated by method 1-3 and the typical occupancy profile is shown in Table 4,5. The simulation using only transition probability matrix is written uncertainty as ”Matrix”.

Table 4. Average temporary concordance rates with the occupancy profile and the generated occupancy patterns.

	Method		
Uncertainty	1	2	3
0%	95.1	95.1	95.1
5%	95.9	94.6	95.6
10%	95.4	94.0	93.6
15%	93.2	94.8	94.4
20%	90.9	93.4	91.1
25%	86.2	93.6	91.5
30%	83.1	92.5	89.6
35%	78.7	90.7	87.5
40%	76.2	86.9	83.3
45%	73.8	83.1	80.4
50%	72.3	79.6	77.2
Matrix	53.5	46.7	44.6

Table 5. Average transitional concordance rates with the occupancy profile and the generated occupancy patterns.

	Method		
Uncertainty	1	2	3
0%	56.8	56.8	56.8
5%	61.4	55.8	59.1
10%	55.1	57.8	52.2
15%	44.0	53.9	52.3
20%	40.0	49.2	42.7
25%	32.4	47.9	35.5
30%	31.5	43.1	34.2
35%	30.9	41.0	32.3
40%	29.1	35.0	30.0
45%	27.4	32.5	26.4
50%	27.6	30.1	27.4
Matrix	21.4	27.8	32.4

The temporary concordance rate of the observed data with the occupancy profile was 60-80%. When the uncertainty factor was specified to a value from 35% to 50% in the case of Method 1 and 50% in the cases of Method 2 and 3, the concordance rates of the generated occupancy patterns was similar to that of the observed data.

The transitional concordance rate of the observed data with the occupancy profile was 30-55%. It can be said that there is a room for improvement on the simulator as the generated occupancy patterns look appropriate from the viewpoint of the transitional concordance while it does not look appropriate from the viewpoint of the temporary concordance rate.

4.3 Multiple Occupancy Profiles

In order to improve the quality of the generated occupancy pattern, an extension to the simulator is proposed in this subsection. Since there might be not only a single occupancy pattern but also multiple patterns, it is possible to achieve more realistic occupancy simulation by employing multiple occupancy profiles. This can be appropriate if one thinks the multiple life patterns such as weekdays and weekends. That is, when the activity of the occupant is generated by the simulator for a long time, the occupant may behave differently based on not only time but also day. In addition, the range of the values of the concordance rates of the occupancy profile with the observed data can be reduced by the use of multiple occupancy profiles depending on the day.

In the computational experiments in this subsection, the transition probability is set to 90%. Each concordance rate is calculated for eight days with each other and reduced into two dimensional vector using principal component analysis to Fig. 2. From Fig. 2, multiple clusters can be recognized. They are manually identified in this paper and for each cluster, the transition matrix is calculated as well as the occupancy profiles. The occupancy profiles are called A, B, and C in this paper.

Table 6 shows the concordance rates of the observed data with the occupancy profiles A, B, and C. The concordance rates of the observed data with the applicable day and these patterns become higher than the first experimental settings. From these results, one can say that the characteristic of multiple occupancy patterns was successfully captured. In addition, Table 7 shows the average concordance rates of the simulation data with the eight days. Since the concordance rates of the generated occupancy patterns was almost similar to that of the observed data with occupancy profiles. From the results in pattern B, C, as for the simulation data generated, it can be seen that the simulator has generated the occupancy patterns more appropriately than the case of only a single occupancy profile. The average concordance rates of the generated occupancy patterns in the case of the single typical occupancy profile obtained from the eight days was 71.2% in Method 1, and 71.2% in Method 2, and 67.2% in Method 3. Occupancy profile B produced the highest concordance rate among the three occupancy profiles for the three methods.

5 Conclusions

In this paper, an occupancy simulator was developed. It was assumed in the simulator that a virtual elderly occupant moves around in a house. The simulator generated an occupancy pattern based on an occupancy profile that is a typical transition between places along with the time information. The occupancy profile was obtained from a set of real data that are obtained through the monitoring of an elderly person who lives alone at a house.

Concordance rates of the observed data with

Table 6. Concordance rates of the observed data with Occupancy profiles A, B, and C (%).

	First	Second	Third	Forth	Fifth	Sixth	Seventh	Eighth
A	100.0	22.6	21.0	37.0	30.1	25.2	34.0	30.5
B	22.2	72.6	67.4	33.5	27.8	39.0	35.3	41.0
C	32.6	35.5	38.4	67.6	69.2	60.0	76.8	67.4

Table 7. Concordance rates of the occupancy patterns generated by the simulator with Occupancy patterns A, B, and C (%).

	A			B			C		
	Meth. 1	Meth. 2	Meth. 3	Meth. 1	Meth. 2	Meth. 3	Meth. 1	Meth. 2	Method 3
A	63.4	66.1	65.8	22.0	22.0	21.7	37.6	35.0	35.1
B	19.5	19.9	19.6	73.6	74.3	74.3	39.8	39.5	39.9
C	29.8	31.2	31.1	37.3	38.8	38.8	62.6	67.8	67.6

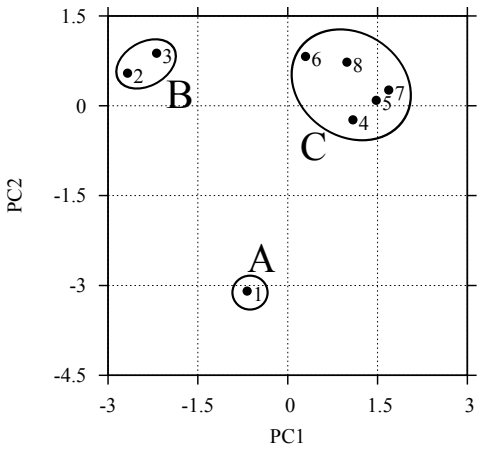


Figure 2. Result of the clustering after reducing the observed data into two-dimansional vector using principal component analysis.

the occupancy profile and that with the simulation data were calculated in order to examine the validity of the generated occupancy patterns by the simulator.

Then, the observed data were grouped though a pre-processing procedure to obtain three occupancy profiles. Another computational experiments were executed to generate occupancy patterns using the multiple occupancy profiles. The validity of the generated occupancy patterns were examined again, and it was shown that these occupancy patterns are more appropriate as simulated occupancy patterns than those generated using only a single occupancy profile.

Future works include the suggestion of the simulation method in order to make the occupancy simulation closer to the real environment.

References

[1] M. J. Akhlaghinia, A. Lotfi, C. Lanensiepen and N. Sherkat, "Single-occupancy simulator for ambient intelligent environment," *Journal of Internet Technology*, No. 9:4, 333 – 338, 2008.

Extracting Descriptive Life Profiles from Mobile GPS Data

Teerayut Horanont† Apichon Witayangkurn† Yoshihide Sekimoto‡ Ryosuke Shibasaki‡

† Institute of Industrial Science, the University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
‡ Cennter for Spatial Information Science, the University of Tokyo
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568, Japan

{teerayut, apichon}@iis.u-tokyo.ac.jp, {sekimoto, shiba}@csis.u-tokyo.ac.jp

Abstract

The pervasiveness of mobile location-based devices that record position, time, and environmental properties has opened new prospects for human mobility research. In this paper, we present an approach to identify user types and location profiles using footprints from anonymous mobile GPS data. The results show promising potential for mining large-scale mobile phone trajectory data to solve real-time urban problems and better understand the complexity of individuals in the urban system.

Keywords: Mobile sensing, GPS, People mobility, Spatio-temporal clustering, Urban planning

1 Introduction

The penetration of the mobile phone exceeds other forms of Information and Communications Technology (ICT) and we are moving closer to having almost as many mobile-cellular subscriptions as people on earth [1]. The mobile revolution is empowering people and transforming our culture into an information society. Along with the advance and spread of mobile sensors, GPS, and their applications, mobile datasets are becoming an incredible source to understand how people behave and use the city.

The ubiquitous GPS location obtained from typical mobile applications provides at least three valid parameters: latitude, longitude, and timestamp. One of the biggest challenges in mining this kind of spatio-temporal data is moving toward the area of big-data analytics that enable us to discover full-scale life patterns in the real world. It is possible to define a higher level of human abstraction by grouping the data into meaningful clusters, according to its similarity. Drawing on unique longitudinal data in a certain period can result in discovering distinct location profiles of individuals while the data generated from mobile devices allows us to empirically examine observable features of a person's movement and thus classify people by type.

There are few previous studies that give the overview of spatio-temporal clustering methods of mobile phone data [2][3]. We can use the mobile phone dataset for various urban applications including estimated daily travel distance [4], analysis of human behavior change during crisis [5], urban monitoring [6], and transportation [7].

A previous study did suggest, surprisingly, that there is a strong regularity in people's mobility which makes 93% of their whereabouts predictable [8]. Most of the recent works discussed the analytical methods using mobile Call Detail Records (CDR) and have not studied a complete set of real world data, especially on a country-wide scale. In contrast to previous studies, this paper investigates a framework to classify user type and extract location profiles from anonymous mobile GPS data.

The rest of the paper is organized as follows: Section 2 describes the dataset used in this research; Section 3 discusses the methodology to uncover the life profiles from mobile GPS; Section 4 provides the discussions.

2 Dataset

2.1 Mobile GPS from mobile location-based application

The dataset used is comprised of the GPS location traces of mobile phone users in Japan. It was collected from users who registered for location-based services from August 1st, 2010 to July 31st, 2011. Technically, a mobile GPS-enabled handset position is measured within five minutes and sent through a network of registered services and analysis engines (Figure 1). To reduce battery consumption, the accelerometer is used to detect periods of relative stasis during which power-consuming GPS acquisition functions can be suspended. Mobile GPS data is provided in a completely anonymous form to ensure privacy of personal information. In total, approximately 9.2 billion records were collected from more than one million mobile GPS users for this analysis. To the best of our knowledge, this is, thus far, the largest number of GPS records collected via a typical cellular operation.

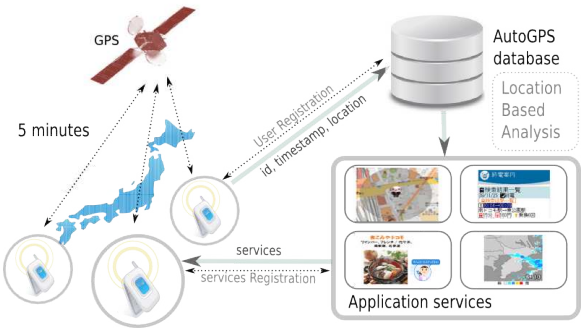


Figure 1. A framework of mobile GPS data collection

2.2 Characteristics of mobile GPS data

Each entry in the dataset included: unique user ID, position (latitude, longitude), timestamp, altitude, and approximated error (i.e., <100 m, <200 m, or <300 m). The average daily GPS points of this dataset is 37, indicating that users in Japan spent approximately three hours traveling with their handset each day. Table 1 shows a sample of collected GPS data while Figure 2 illustrates a one-day GPS trace for an ordinary user in the Tokyo area.

Figure 2. A sample of one-day GPS trace

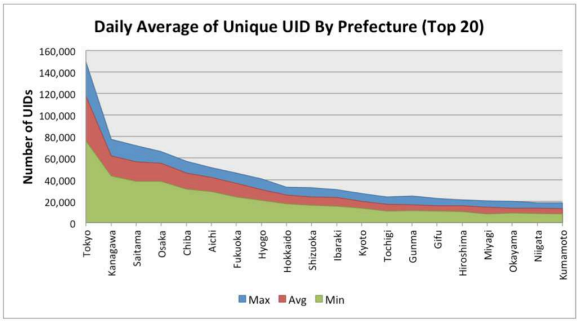


Figure 3. Daily average count of unique mobile GPS users by prefecture

Table 1. Sample of mobile GPS data

Unique ID	Time	Latitude	Longitude	Error	Altitude
0000001	2010-08-0122:24:35	35.362259	136.812857	3	51
0000001	2010-08-0101:34:53	35.362748	136.814160	1	64
0000005	2010-08-0122:45:07	33.278713	130.511025	3	59.86
0000005	2010-08-0122:40:08	33.245577	130.508839	3	47.08
0000005	2010-08-0122:25:12	33.210181	130.504812	2	64.02
0000005	2010-08-0121:35:30	33.306663	130.512514	1	64
0000005	2010-08-0121:23:35	33.389261	130.551836	2	51.58
0000005	2010-08-0118:58:10	33.584711	130.397758	1	64
0000005	2010-08-0118:33:10	33.587222	130.401943	3	64

2.3 Limitations

Currently, the Mobile GPS dataset reflects about only 1.3% of the Japanese population. Figure 3 shows the daily average count of unique IDs by prefecture during a one year acquisition period. There is an uncertainty of period of use since subscribers can suspend a mobile GPS service at any time using a simple online application. Not every user is expected to continue GPS tracking over the entirety of their service.

3 Uncover Life Profile using Mobile GPS data

The usage of real-life spatio-temporal data to extract a comprehensive description of individuals and urban metabolism is becoming increasingly important in today's society. The clustering technique is one of the fundamental approaches to reduce the complexity and discover similarities in large datasets. In this section, we describe the technique used to classify people type and identify their life profiles from the entire mobile GPS trajectories. We developed methodologies and a spatio-temporal analysis platform based on Hadoop [9] to mine the data and discover the particular characteristics of a certain group of people.

3.1 Deriving of activity profile

For decision-making and information processing to understand user behavior under specific conditions, user type, trip purpose, and geographic area are the important parameters. In contrast to a traditional survey, without any previous knowledge of the personal features of the mobile phone users, this mobile GPS data may not be useful. Since mobile GPS acquisition was toggled by an accelerometer, the GPS points were created only when the users moved with their devices. High movement in a specific time period and activity in a specific geographical space where the user stays can be inferred as a life profile. For instance, the long absence of GPS points in the night may be inferred a home location, and high GPS acquisition in the morning and evening may be inferred as the commute from home to work or work to home. Since the individual movement can be clearly captured, the daily profile can be observed by hourly grouping of the GPS's active periods.

To cluster people into groups, we used the Iterative K-Means method to classify the user by similarity of movement activity. Cosine similarity is used as the distance function for clustering and Bayesian Information Criterion (BIC) is used for cluster evaluation to identify the optimum number of result clusters. Figure 4 describes the framework to calculate and obtain the activity profile from the mobile GPS dataset.

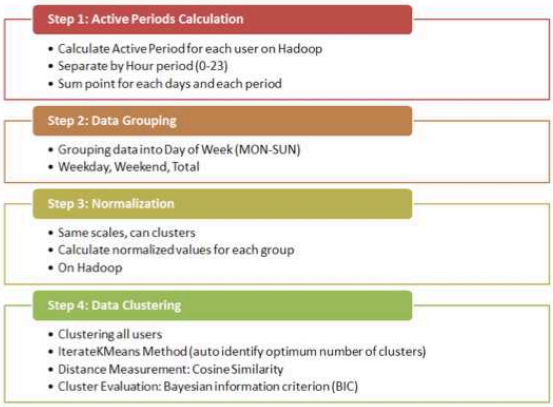
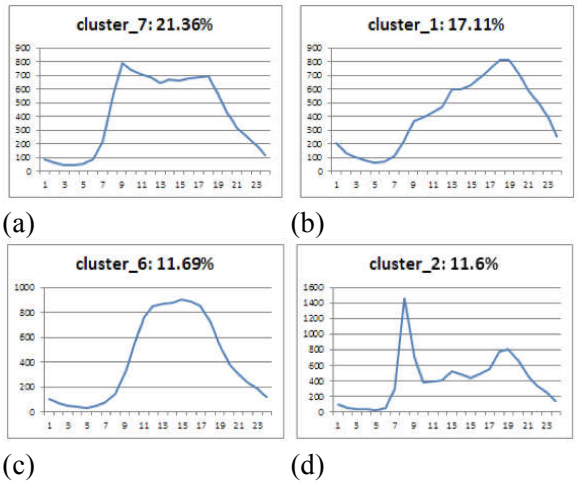
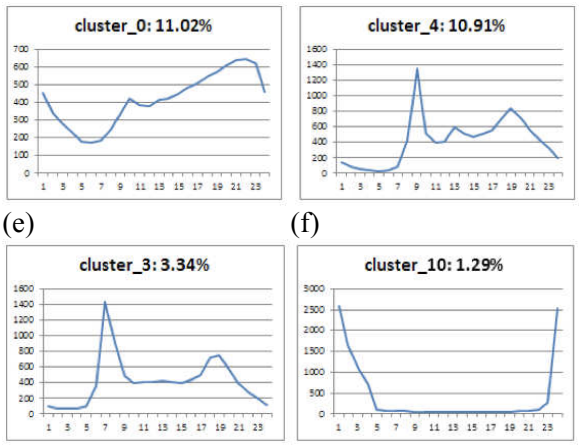


Figure 4. Calculation framework of activity profile deriving process.

To reduce bias in this calculation, weekends and holidays for the entire calendar year have been discarded from the dataset. As a result, the one-year activity of 1.5 million mobile GPS users has been classified into 24 groups. We decided to use the first eight clusters to do further analysis since they represent more than 88% of the entire population. Figure 5 illustrates the life profile of the eight major clusters.





(e) (f)

(g) (h)

Figure 5. The activity profile of possible (a) ordinary daytime worker; (b) high evening activity worker; (c) high midday activity worker; (d) ~8:00 AM commuting time office worker; (e) late night high activity worker; (f) ~9:00 AM commuting time office worker; (g) ~7:00 AM commuting time office worker; (h) night time worker

From the activity profiles, it is clear that the majority of people are similar in starting their activity in the morning, have high activity during the day and lower activity at night. Two commuting periods, to and from home and work place, can be clearly captured, as seen in the graph as morning peak and evening peak. Morning peaks come with a spike indicating that people tend to go out to work at the same time. However, we can observe three slight shifts of morning peak from three different clusters (d, f, and g). This is because the starting work hour is different for different business entities. On the other hand, evening peak has a wider range of active movement. It indicates that people do more activities in the evening and tend not to go back home at the same time.

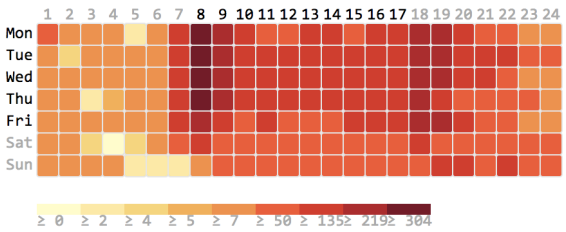


Figure 6(a)

The individual activity profiles can be constructed in different temporal dimensions based on aggregate statistics. (Figure 6)

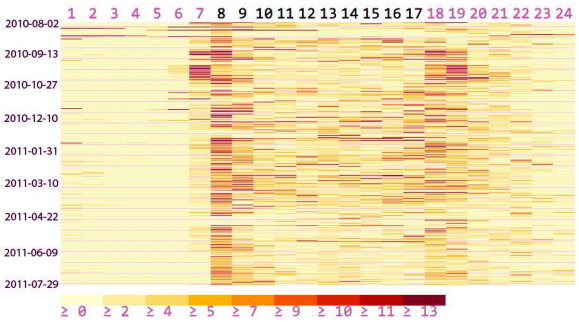


Figure 6(b)

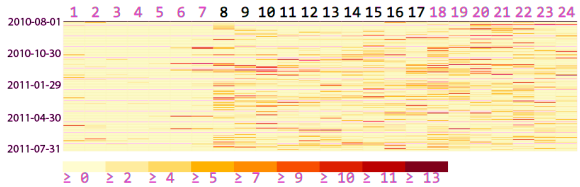


Figure 6(c)

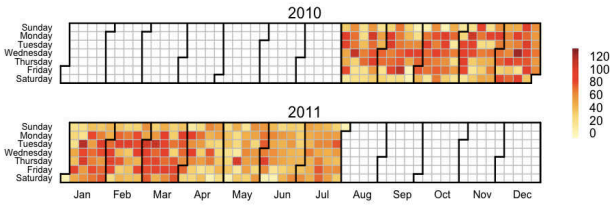


Figure 6(d)

Figure 6(a) represents weekly activity profiles of one sample in cluster 2; Figure 6(b) shows accumulated annual profiles highlighting their activity by weekday stack (245 days); Figure 6(c) illustrates annual profiles highlighting their activity by weekend and holidays stack (120 days) and Figure 6(d) shows the overall annual activity profiles aggregated by day.

These patterns can be used for a detection approach based on the discovery of correlated abnormal pattern. For example, if there is a sharp increase or decrease in an activity profile, it is highly likely that there is an anomaly event appearing in the correlated time slot.

3.2 Deriving space profiles

People are inherently instilled with routine across all temporal scales, from daily actions to monthly or yearly patterns. They usually spend most of their time at a few key locations, such as

home, work, and some favorite place. Being able to identify individual space profiles is crucial for effective city planning and public infrastructure management. First, we have identified stay points as sequences of consecutive GPS records. From our previous experiment [5], we define a stay point as a GPS mobile interval of twenty (20) minutes minimum time and a maximum coverage distance of 150 meters (Figure 7). The output stay points are then combined together and used as an input for a significant place finding algorithm.

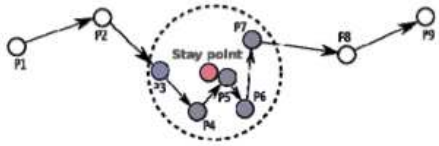


Figure 7. Stay point extraction method.

The DBSCAN clustering technique is used for this purpose. This density-based clustering method is efficient in eliminating outliers within the cluster. We validated the significant place finding with census data provided by the Statistics Bureau, Ministry of Internal Affairs and Communications. The estimation of finding a home location is good with a resulting correlation coefficient, $R^2 = 0.84$.

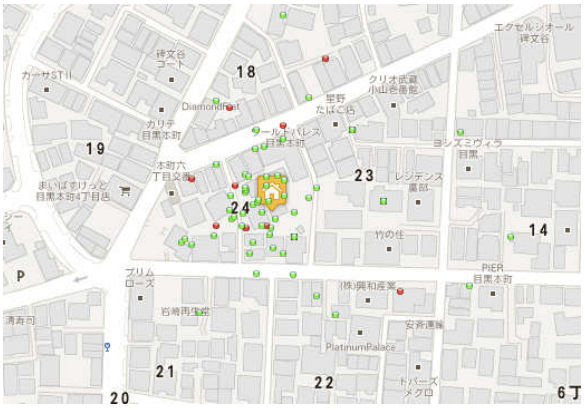
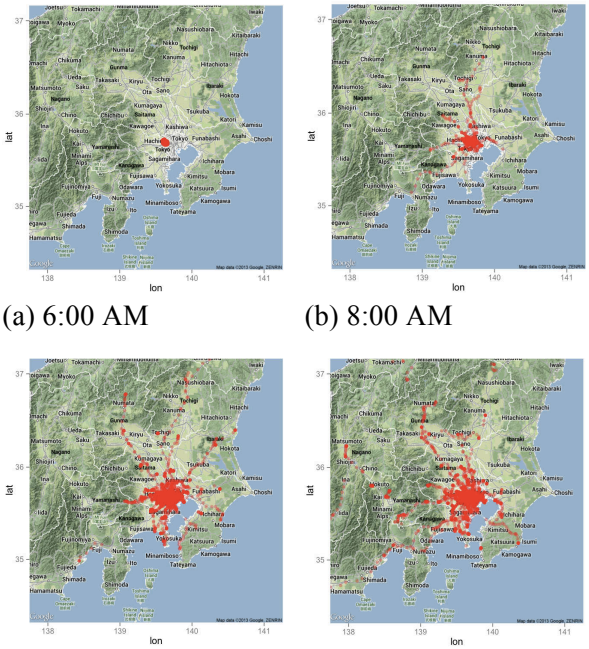


Figure 8. Estimation of home location.

Figure 8 illustrates the predicted home location from our spatio-temporal visualization and analysis platform. The red dots are daily anonymous mobile GPS points; the green dots are stay points calculated by our platform using one month of data. By using the proposed clustering algorithm in combination with a multiple criteria decision tree, home location and space profiles of individuals are created.

Illustration of urban metabolism. The extracted stay points can be conditional on the places that people are active during the day. Associating time-series stay points to the city space gave us an image of urban metabolism evolution. Figure 9 illustrates the idea of urban metabolism of one municipal in Tokyo. The analysis reveals a concept of urban metabolism in order to understand the complex, and often invisible, set of interactions between humans and the city. The flow of people can be tracked as a transformation of mass and energy within an urbanized sub-region over time periods. This time-series information provides valuable knowledge to decision makers responsible for city management. This is especially important for transportation planning, event management, and disaster preparedness.



(a) 6:00 AM (b) 8:00 AM

(c) 12:00 AM (d) 4:00 PM

Figure 9. Explanation of the coverage and distance traveled by people of a sample municipal in Tokyo.

4 Discussions

The popularity of the smart phone made location information widely available. This study has investigated design solutions for extracting life profiles from anonymous mobile phone GPS

data. Our methodological framework makes the entire city's activity visible and provides the ability to identify daily commute distances and the common places for individuals and groups in a more dynamic manner.

Mining of this mobile GPS data offers the opportunity to model human behavior and social life at an unprecedented scale and fidelity. Our findings confirm results from previous studies indicating that people have very much alike in daily routines and that they can be characterized by a few representative groups. In the case of Japan, the extracted life profiles of eight major classes give us the opportunity to model nearly 90% of the entire population. Considering the temporal dimension, we found daytime workers are 88% dominant compared with 12% for evening and nighttime workers. We also found that no specific group is very highly dominant. People have similar daily routines; they go to work and return home as indicated by the morning peak and evening peak in most of the clusters. Although the different starting work time can be clearly recognized by the system, the algorithm still lacks the ability to recognize occupation types. This will need to be addressed in future works.

Grouping similar life profiles is a fundamental step for location-based applications such as marketing strategies. The findings can potentially suggest specific customer areas for an analyst to target. It is also useful for a city to use patterns that are populated by real time information and model projections to maintain and update city management.

Finally, we confirm our approach of using mobile phone and information technology in real time city management practices. The results give some perspectives in analyzing spatially-rich mobility data that can be used to explain and predict a range of important urban phenomena.

Acknowledgment

This work was supported by the DIAS (Data Integration and Analysis System) and GRENE (Green Network of Excellence) project of MEXT (Ministry of Education, Culture, Sports, Science and Technology). Special thanks to ZENRIN DataCom CO., LTD for archiving the dataset used in this research.

References

[1] ICT Facts and Figures 2013, International Telecommunication Union (ITU), press release pp 1-8, 2013.
[2] Y. Yuan and M. Raubal, "Spatio-temporal knowledge discovery from georeferenced mobile phone data," pp. 121–126, 2010.
[3] N. Andrienko and G. Andrienko, "A visual analytics framework for spatio-temporal analysis and modelling," Data Mining and Knowledge Discovery, pp. 55–83, Aug. 2012.
[4] S. Isaacman, R. Becker, R. Cáceres, S. Kobourov, J. Rowland, and A. Varshavsky, "A tale of two cities," Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications - HotMobile '10, p. 19, 2010.
[5] T. Horanont, A. Witayangkurn, Y. Sekimoto, R. Shibasaki, "Large-Scale Auto-GPS Analysis for Discerning Behavior Change during Crisis," IEEE Intelligent Systems, 11 Jan. 2013. IEEE computer Society Digital Library.
[6] T. Horanont and R. Shibasaki, "An Imple-mentation of Mobile Sensing for Large-Scale Urban Monitoring," International Workshop on Urban Community and Social Applications of Networked Sensing Systems, pp. 51–55, 2008.
[7] O. Järv, R. Ahas, E. Saluveer, B. Derudder, and F. Witlox, "Mobile phones in a traffic flow: a geo-graphical perspective to evening rush hour traffic analysis using call detail records.," PloS one, vol. 7, no. 11, p. e49171, Jan. 2012.
[8] M. C. González, C. a Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns.," Nature, vol. 453, no. 7196, pp. 779–82, Jun. 2008.
[9] A. Witayangkurn, T. Horanont, and R. Shibasaki, "Performance comparisons of spatial data processing techniques for a large scale mobile phone dataset," Proceedings of the 3rd International Conference on Computing for Geospatial Research and Applications - COM.Geo '12, p. 1, 2012.

Towards the Use of Upper Ontologies for Refugee Emergencies in Disaster Management

Pasinee Apisakmontri[†*], Ekawit Nantajeewarawat[†], Marut Buranarach[‡], and Mitsuru Ikeda[*]
[†]School of Information, Computer, and Communication Technology,
Sirindhorn International Institute of Technology, Thammasat University, Thailand
[‡]National Electronics and Computer Technology Center, Thailand
[*]School of Knowledge Science, Japan Advanced Institute of Science and Technology, Japan
pasinee.a@jaist.ac.jp, ekawit@siit.tu.ac.th,
marut.buranarach@nectec.or.th, ikeda@jaist.ac.jp

Abstract

Disaster information, such as information on earthquakes, tsunamis, floods, fire, hazard chemicals, hurricanes, etc., is usually rapidly changeable, ambiguous, and huge. It often involves data items from diverse domains on the web, some of which have similar meanings but appear structurally dissimilar in different data sources. A promising remedy is to provide a basis for common understanding of disaster management in a formally correct way. Formal upper ontologies provide rich vocabularies for describing an information system, and they form a basis for defining abstract data types that support interoperation among heterogeneous systems. Development of an ontology for Humanitarian Aid for Refugee in Emergencies (HARE) is presented throughout this work.

Keywords: Disaster Management, Ontology, Interoperability, Humanitarian Aid for Refugee in Emergencies, Upper ontology

1 Introduction

Success of disaster information management depends on finding and effectively integrating related information to take decisions during information distribution [1]. A survey in [2] reported that ontologies and current technologies can be used to identify and associate semantically corresponding concepts in the related disaster information in order that heterogeneous data can be integrated. An ontology is collections of information that provided by the Semantic Web technology. It gives advantages to determine common vocabularies for knowledge representation and specifies constraints on the relationships between objects in an application domain. Such advantages cause information to be given well-defined meaning and better enabling computers and people to work in cooperation [3].

The term disaster management is used to describe large-scale processes of disaster relief and disaster recovery. It may refer to an extension of local emergency management. Disaster management is typically not an isolated process. It must be quickly aware of information from many sources and also must feed information externally [4]. Nowadays, there are many software applications launched to facilitate disaster management tasks. Combining information from various sources in these applications is, however, often problematic. Semantic interoperability of disaster information is important in these applications, since they were not initially developed for integrating with each other. Semantic interoperability allows the applications to cooperate with minimal modifications [5].

In this paper, we focus our attention on development of an ontology for Humanitarian Aid for Refugee in Emergencies (HARE), which is an important part of the disaster management domain. The paper is organized as follows: Section 2 provides background and related works. Section 3 describes our ontology development methodology. Section 4 concludes the paper.

2 Background and Related Works

2.1 Ontologies

An ontology is a conceptualization of data domain, which has the formal explicit definition of concepts (known as classes) in the given domain(s), properties (known as attributes) describing various features of the concepts, restrictions on these properties and well-defined relationships between concepts. The instances to represent the concept of ontology are known as

Individuals for the formal concept definitions, constitutes a knowledge-base. The target of ontology in Information Science is to define a common vocabulary and describe words and their interrelationships on a formal and hierarchical level. In addition, the ontology is also to enable knowledge sharing and data consistency [6].

2.1.1 Ontologies and Disaster management

Disaster management will have the success of management as getting the right resources to the right place at the right time; to provide the right information to the right people to make the right decisions at the right level at the right time. Interoperability is one challenging when involving organizations try to integrate these individual data sources [7; 8; 9].

The three types of interoperability are system interoperability, syntax and structure interoperability and semantic interoperability. During disaster management, they also have problems with interoperability at a system, a syntax and a semantic level. A system and syntax level has been achieved by hardware improvement and a syntax standard respectively. For semantic interoperability, a standard terminology is dependent upon the context of use and users [10]. The use of different terms and approaches cause confusion in the specification of universally accepted entities, concepts, rules, relations, and semantics. The basis of an ontology can cope with the confusion because an ontology is a logical theory accounting for the intended meaning of a formal vocabulary [1; 11].

Therefore, the disaster management domain is still a lack of a terminology and relationship among other data sources. From those causes, the semantics of disaster information are unclear for integrating systems, so that the use of ontologies is to better allow information interoperability and information sharing to exchange their information immediately in a short time period. Ontologies are similar to a conceptual schemata in database systems, but ontology provides a logical description to allow application systems to integrate together without having to share data structures. Because of the potential of an ontology, one data source can be combined with another. Thus, the need for ontologies to establish a common specification to deal with the entire disaster management cycle is necessary.

2.1.2 Upper ontology

An upper ontology can be also called as a top-level ontology or foundation ontology, which describes the general concepts for all knowledge domains. As general concepts, several ontologies can be hierarchical implementation on the upper ontologies. Although upper ontology has several advantages, such as integrating with existing ontologies, providing a predefined set of ontological entities, turning away from conceptual ambiguity, however, semantic interoperability between ontologies is the main advantage of upper ontology [12; 13; 14]. Based on reviews of the characteristics of ontologies to be used as upper ontology, we decide that the below ontologies meet most of our requirements.

2.1.3 Related upper ontologies

- **Semantic Web for Earth and Environmental Terminology (SWEET)** The SWEET ontologies include several thousand terms, spanning a broad extent of Earth system science and related concepts (such as data characteristics). To support such a large collection and adhere to the guiding principles, the concepts are divided, where possible, into orthogonal dimensions or facets in support of reductionism [15].

- **Suggested Upper Merged Ontology (SUMO)** It is the most outstanding proposal under consideration by the IEEE Standard Upper Ontology (SUO) working group. Its effort is to link categories and relations that come from different top level ontologies in order to improve ontologies in the Semantic Web area. SUMO is based on the fundamental distinction between Physical and Abstract. The goal of SUMO is to develop a standard ontology that promotes data interoperability, information search and retrieval, automated inference, and natural language processing. SUMO is implemented in the first-order logic language SUO-KIF that can be automatically translated into OWL, although the translation is lossy. The ontologies that extend SUMO are available under GNU General Public License [16].

- **Friend of a friend (FOAF)** An ontology which used to describe people, their ac-

tivities and their relations to other people and objects. The FOAF is useful in social networks with decentralized technology for connecting public web sites [17].

- **Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)** DOLCE belongs to the WonderWeb library of foundational ontologies that provide a set of upper level concepts. It aims at capturing the ontological categories, and underlying natural language and human common sense that assist in making formed conceptualizations explicit. Furthermore, a combination of DOLCE has been published to improve some parts of DOLCE. DOLCE is based on the fundamental distinction between Endurants (i.e., objects or substances) and Perdurants (i.e., events or processes) [18].

2.2 Humanitarian Aid for Refugee in Emergencies (HARE)

Disaster management has four phases, which involve mitigation, preparedness, response, and recovery when natural or human-made disaster occurs [19]. The mass of information is spread over countries in crisis. It is necessary to develop information systems to support that huge information. An IT infrastructure must be established as an information controller in each system. The HARE is one component that emphasizes the response and recovery phases. The objectification of HARE is under an obligation to lead and coordinate international action from the world-wide protection of refugees and manage the resolution of refugee problems. This paper presents a novel concept under the UNHCR operations [20]. The UNHCR stated that a refugee emergency of UNHCR is any situation in which the life or well-being of refugees will be threatened unless immediate and appropriate action is taken, which demands an extraordinary response and exceptional measures. The HARE has been cooperating with many systems both domestic and international. Obviously, information systems are critical for the HARE. In the past, various systems of the HARE were focused on a stand-alone system, but currently efficient systems should rely on other systems. Thus, the interoperability of systems is challenging. This paper tries to establish semantic interoperability of heterogeneous data from several operations of UNHCR.

3 Methodology

3.1 Ontology building methodology

Ontologies provide semantic interoperability between people and systems by providing explicit specification of conceptualizations. Formal specification reduces an ambiguity in definitions. People generally use ontologies to express models of the problem domain. They also use ontology as a knowledge base to either extract new knowledge, or support several functions of their systems. In this paper, we apply the ontology for HARE domain from UNHCR's subsystems, namely ,Refugee Registration System, Identification of Persons of Concern System, Emergency Planning System, Distribution of Assistance System, and Donation System to support the semantic interoperability . The establishment of an ontology for the HARE domain is a pre-requisite to developing an application. There are several methods to building ontologies. Most of the methods for building ontologies are focused on the conceptualization and ontology implementation, but they do not pay enough attention for integration of the ontologies [10; 21]. We adopt the following three steps from Uschold and Kings method [22]:

1. Identify the purpose: to consider the requirement of Refugee Emergencies ontology and create the process models

2. Ontology capture: to consider knowledge models from the process models

3. Coding and integrating: to implement the Refugee Emergencies Ontology and integrate with upper ontologies

3.1.1 Identify purpose

In order to establish the ontology requirements, we have read the overall operations of the HARE from the literature on Handbook for Emergencies [20] to undertake the abstraction and processes of refugee emergencies from UNHCR, and we learnt that the HARE has responsibilities for resolution of refugee problems. The operations of UNHCR cover many areas in refugee emergencies, including health, food,

sanitation and water, as well as key field activities corroborate the operations such as logistics, community services, and registration. Such operations must be managed and controlled by many associate organizations. In the identify purpose step, information should be extracted carefully from documentations. Misunderstanding in this step could lead to misinterpretations in next steps. We determined the domain, scope and purpose of the operations into the refugee emergencies process models that depict the principle of refugee emergencies' operations.

Process 1: Refugee Registration This process is capturing the implemented process of the refugee registration in an emergency process model. Refugees should be registered as fast as possible after reach to a refugee center. Refugee profile must be the first information that organizations would like to know. The following information is recorded for a person of concern individual verification: Name, Unique identifying registration number, Date and place of birth, Sex, Existing identity documents, Marital status, Special protection and assistance needs, Level of education, Occupational skills, Religion, Language, Household and family composition, Date of arrival, Current location and address, Place of origin, Photograph. This information will be collected to be the properties of Refugee profile concept.

Process 2: Identification of Persons of Concern System After the refugee registration process, if time permits, a pre-screening should take place at this stage to identify those who may not be of concern to UNHCR. The refugee profile will be analyzed for refugees needs assessment. An accurate estimate of numbers of refugees is a prerequisite for effective protection and assistance, and identification of beneficiaries, including persons with special needs.

Process 3: Emergency Planning System The planning process is very important. Efforts should be made to design and implement a shelter as soon as possible. Several organizations must rely on this planning system. The project is the structural planning for such as the training, logistics, telecom, security, sites(camp, shelter), etc. This will enable better management such as shelter management, non-food and food items' distribution.

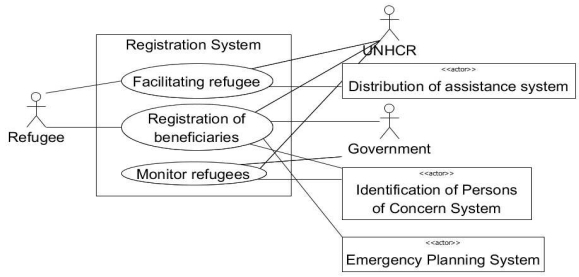


Figure 1. Registration System Diagram

Process 4: Distribution of Assistance There is a simple system to handle the distribution of assistance and provision of service to refugees, including emergency health care, distribution of food and non-food items. Many staffs from organizations participate in this system for sharing help to refugees.

Process 5: Donation System The Donation system is designed to receive, manage, and distribute a mass of donated goods and services. With the help of refugee communities, they identify refugee individuals and groups with their needs, especially unaccompanied and separated children. This system operates with the distribution through an assistance system to support needs of refugees.

3.1.2 Ontology capture

In the second step, the refugee emergencies process models from the previous step will set the initiation to establish ontological conceptualization, which is the key to the HARE ontologies. In this step, we developed the Unified Modeling Language (UML) Use Case Diagram that is used to create the functionality for the system to graphically represent and envision the concepts and relationships between elements. This paper selects five principles of the HARE in order to understand an extent to which interoperability can be supported. (Fig.1-5)

3.1.3 Coding and integrating

In the Refugee Emergencies ontology implementation step, the core concepts and the existing relations from use case diagrams help to build the scope and role of the ontology. Each use case consists of detail to extract relevant classes, their scopes and the required detail. Firstly, a first set of terms is gathered by analyzing use cases. The selected terms are consid-

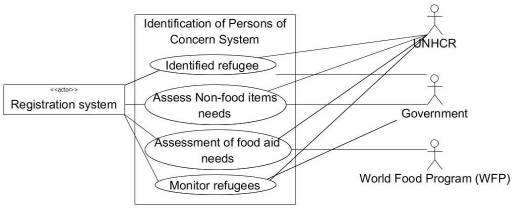


Figure 2. Identification of Persons of Concern System Diagram

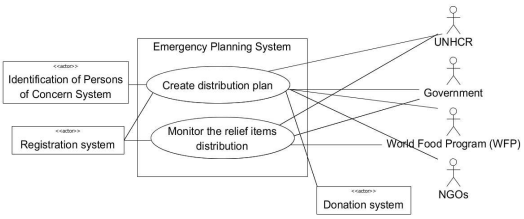


Figure 3. Emergency Planning System Diagram

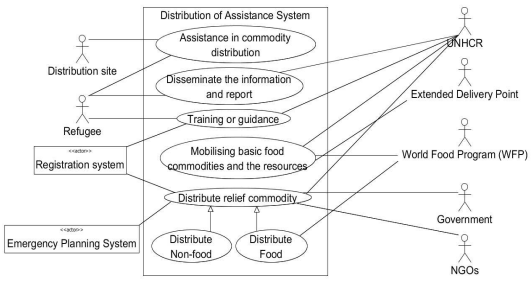


Figure 4. Distribution of Assistance System Diagram

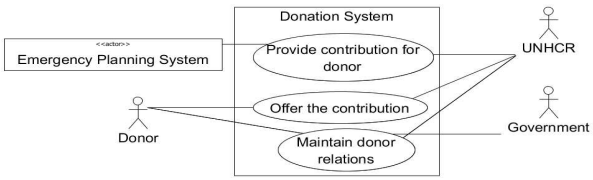


Figure 5. Donation System Diagram

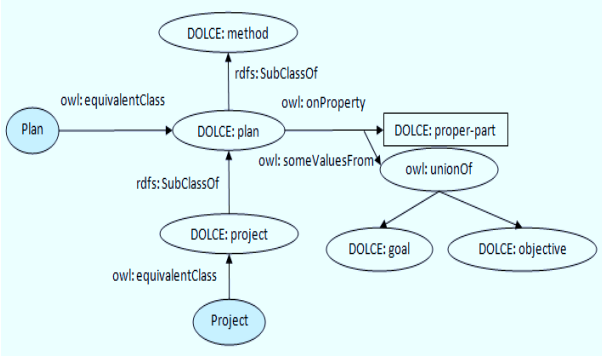


Figure 6. DOLCE upper ontology and HARE ontology

ered as the core concepts of Refugee Emergencies ontologies. That core concepts in the operations of the HARE are Refugee, RefugeeActivity, RefugeeNeed, Commodity, Person, Plan, Project, Organization, Staff. After the core concepts are defined, subclasses and disjoint decompositions are also identified, such as a food product is a particular type of Commodity. A commodity is characterized by Unit. Each commodity should be assigned a unit, which means that a relation "hasUnit" between classes Commodity and Unit must be established (Fig.9). To get core concepts, the relations between concepts and axioms are important to define implicit meanings.

The implementation of the HARE ontology required to choose an appropriate ontology editor and development environment. The Protégé development platform which contains the Protégé-OWL ontology editor for the Semantic Web is chosen for using in this research. After creating ontology, we would notice that there are some classes in the ontology that can be hierarchical implementation on the upper ontologies. As the related upper ontology review, we found that the concepts of the upper ontologies, which are matched with the HARE ontology, are in the DOLCE, SWEET, SUMO, and FOAF. They can provide the basis for common understanding in HARE as shown in Fig.6-10.

We describe the core classes that are derived from each use case with their upper ontologies and relations as follows:

Use case 1: Refugee Registration (Fig.1) The main concepts of the proposed ontology are the Refugee, Person, and RefugeeActivity concept. Those concepts can be defined by the ex-

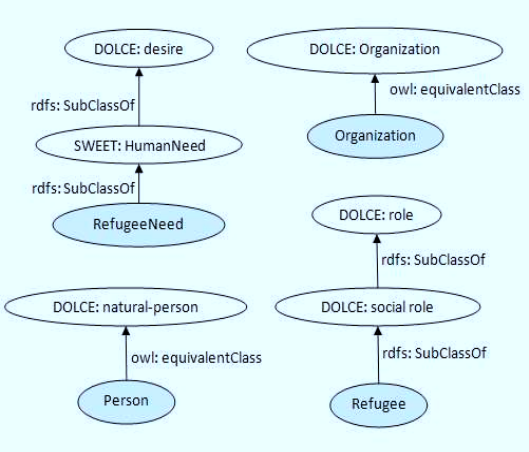


Figure 7. DOLCE upper ontology and HARE ontology

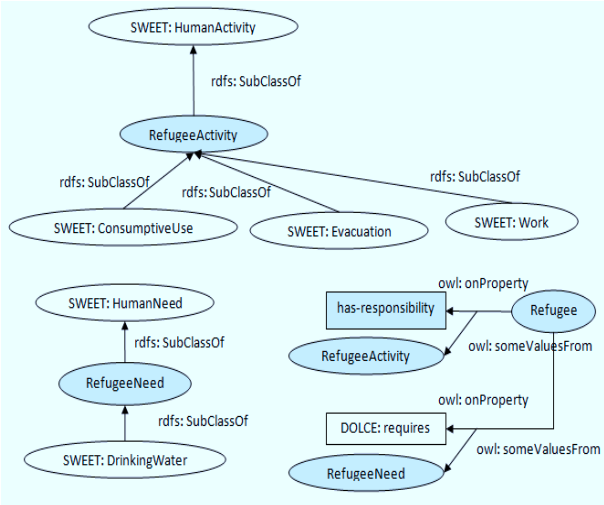


Figure 8. SWEET upper ontology and HARE ontology

isting ontology as follows:

- Refugee is a subclass of DOLCE: social role(Fig.7).
- Refugee has at least one hasResponsibility relationship to individuals that are member of RefugeeActivity (Fig.8).
- RefugeeActivity is a subclass of SWEET: HumanActivity and has SWEET: ConsumptiveUse, SWEET: Evacuation, and SWEET: Work as its subclasses (Fig.8).
- Person is equivalent to DOLCE: natural-person (Fig.7) and an individual of Person has only 1 string to FOAF: familyName, FOAF: firstName, and FOAF: lastName DataProperty(Fig.10).

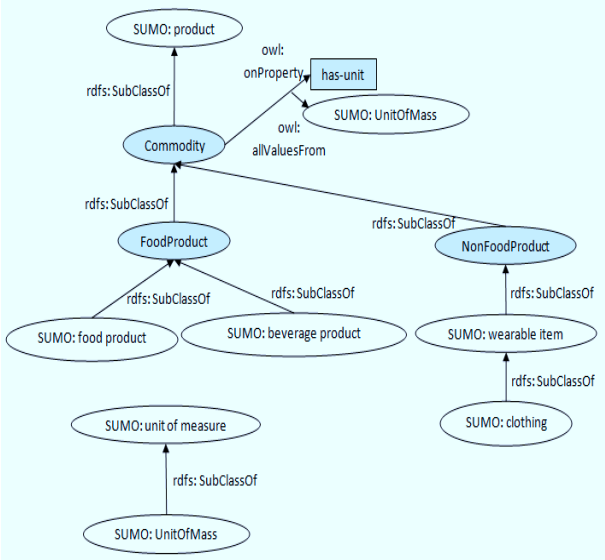


Figure 9. SUMO upper ontology and HARE ontology

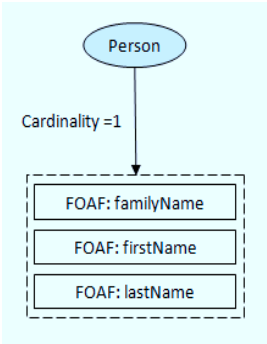


Figure 10. FOAF upper ontology and HARE ontology

Use case 2: Identification of Persons of Concern System (Fig.2) The main concepts of the proposed ontology are Refugee and RefugeeNeed concept. This concept can be defined by the existing ontology as follows:

- Refugee has at least one requires relationship to individuals that are member of RefugeeNeed (Fig.8).
- RefugeeNeed is a subclass of SWEET: HumanNeed and has SWEET: DrinkingWater as its subclass(Fig.8).

Use case 3: Emergency Planning System (Fig.3) The main concepts of the proposed ontology are Plan, Project, and Commodity concepts.

This concept can be defined by the existing ontology as follows:

- Plan is equivalent to DOLCE: plan and DOLCE:plan has at least one DOLCE:proper-part relationship to individuals that are member of DOLCE: goal or DOLCE: objective (Fig.6).
- Project is equivalent to DOLCE: project and DOLCE: project is a subclass of DOLCE: plan (Fig.6).
- Commodity is a subclass of SUMO: product and has food product, and non-food product as its subclasses (Fig.9).
- Commodity has only has-unit relationship to individuals that are member of SUMO: UnitOfMeasure (Fig.9).
- FoodProduct has SUMO: food product and SUMO: beverage product as its subclasses and NonFoodProduct has SUMO: wearable item as its subclass (Fig.9).

Use case 4: Distribution of Assistance (Fig.4) The main concepts of the proposed ontology are Refugee, Commodity, and Organization concept. This concept can be defined by the existing ontology as follows:

- Organization is equivalent to DOLCE: Organization (Fig.7).

Use case 5: Donation System The main concepts of the proposed ontology are Commodity, Refugee, and RefugeeNeed. So that the concepts in this use case are as same as the above concepts that we mentioned.

After we completed all above steps, our HARE ontology has 89 concepts, 17 object properties, 26 data properties, and 82 individuals (Fig.11). The ontology we created can be used to be a common conceptualization of HARE that other related systems would be integrated with in this domain.

4 Conclusion

Semantic interoperability is critical to Disaster Management information sharing and reusing knowledge. In this paper, we proposed the semantic interoperability to the HARE domain base on the UNHCR. We presented the HARE ontologies design choices, and adopted an ontological approach for semantic interoperability to be the common understanding of any information. We presented our ontology design based on an ontology development method and considered the requirement from many documentations that reflect the humanitarian aid system in reality. In addition, after we design the HARE ontology, we provided a comprehensive review of related upper ontologies, which can be integrated with our ontology. The upper ontologies help avoiding ambiguities in the large-scale information integration. Therefore, we believe that the proposed HARE ontologies with upper ontology integration are reasonable for reducing problems on semantic interoperability. The suitable interoperability between the HARE systems can make the efficient access to humanitarian aid during disasters and the ability to save a greater number of lives. Despite being an initial exploratory research, this work is an important step towards exploring in the greater detail and implementing the Semantic Web technologies in the Disaster Management domain.

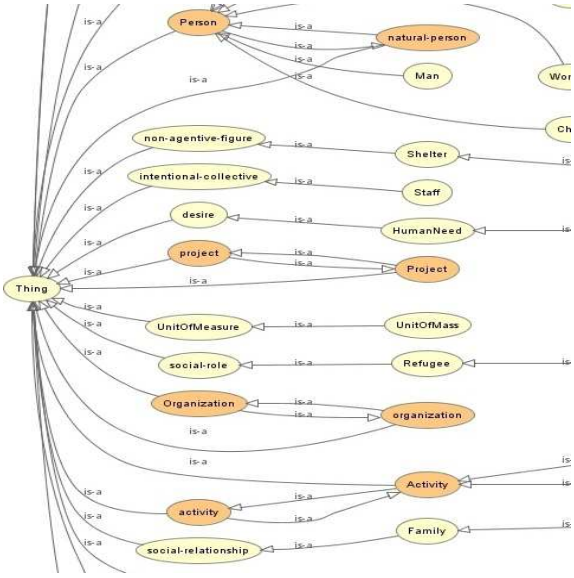


Figure 11. Example of HARE ontology

Acknowledgment

I would like to express my very great appreciation to my research supervisors for their valuable and constructive suggestions during the planning and development of this research work and also thank for excellent information from United Nations High Commissioner for Refugees (UN-

HCR). This research was conducted under a grant in the SIIT-JAIST-NECTEC Dual Doctoral Degree Program.

References

[1] W. Xu and S. Zlatanova. Ontologies for disaster management response. *Geomatics Solutions for Disaster Management*, pages 185–200, 2007.

[2] V. Hristidis, S.C. Chen, T. Li, S. Luis, and Y. Deng. Survey of data management and analysis in disaster situations. *Journal of Systems and Software*, 83(10):1701–1714, 2010.

[3] Eric Miller. W3c technology and society domain: Semantic web points. `www.w3.org/2001/sw/EO/points`. Accessed: 04/09/2012.

[4] T. Berners Lee. Disaster management. `www.w3.org/wiki/DisasterManagement`. Accessed: 04/09/2012.

[5] M. Missikoff and F. Taglino. An ontology-based platform for semantic interoperability. *Handbook on Ontologies*, pages 617–634, 2004.

[6] Wikipedia. Ontology (information science). `http://en.wikipedia.org/wiki/Ontology_(information_science)`. Accessed: 21/01/2013.

[7] KR Ratnam and DD Karunaratne. Application of ontologies in disaster management. 2008.

[8] Mandana Sotoodeh. *Ontology-Based Semantic Interoperability in Emergency Management*. PhD thesis, The University of British Columbia, July 2007.

[9] P. Di Maio. An open ontology for open source emergency response system. *Open Source Research Community*, 2007.

[10] J. Garrido and I. Requena. Proposal of ontology for environmental impact assessment: An application with knowledge mobilization. *Expert Systems with Applications*, 38(3):2462–2472, 2011.

[11] Z. Fan and S. Zlatanova. Exploring ontologies for semantic interoperability of data in emergency response. *Applied Geomatics*, 3(2):109–122, 2011.

[12] A. Gómez-Pérez and M. Fernández-López. *Ontological engineering : with examples from the areas of knowledge management, e-commerce and the semantic web*. Springer, 2010.

[13] Robert Hoehndorf. What is an upper level ontology? `http://ontogenesis.knowledgeblog.org/740`, 2010.

[14] Ajay Tripathi and Hassan A. Babaie. Developing a modular hydrogeology ontology by extending the {SWEET} upper-level ontologies. *Computers Geosciences*, 34(9):1022 – 1033, 2008.

[15] R.G. Raskin and M.J. Pan. Knowledge representation in the semantic web for earth and environmental terminology (sweet). *Computers & Geosciences*, 31(9):1119–1125, 2005.

[16] Suggested upper merged ontology (sumo). `http://www.ontologyportal.org/`. Accessed: 21/05/2013.

[17] The friend of a friend (foaf) project. `http://www.foaf-project.org/`. Accessed: 21/05/2013.

[18] Dolce : a descriptive ontology for linguistic and cognitive engineering. `http://www.loa.istc.cnr.it/DOLCE.html`. Accessed: 21/05/2013.

[19] Renato Iannella. Emergency information interoperability frameworks. `www.w3.org/2005/Incubator/eiif/XGR-Framework-20090806/`. Accessed: 04/09/2012.

[20] The Emergency Preparedness and Response Section. *Handbook for Emergencies*. United Nations High Commissioner for Refugees, UNHCR Headquarters, Case Postale 2500, Switzerland, 3 edition, 2007.

[21] Sebastian Scheuer, Dagmar Haase, and Volker Meyer. Towards a flood risk assessment ontology knowledge integration into a multi-criteria risk assessment approach. *Computers, Environment and Urban Systems*, 37(0):82 – 94, 2013.

[22] Mike Uschold, Michael Gruninger, Mike Uschold, and Michael Gruninger. Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, 11:93–136, 1996.

Line Segmentation for Khmer Printing Text Document

Dina Sok† **Sovila Srun‡*** **Por Heng*** **Beanbonyka Rim**** **Khoy Saovorak***** **Des Phal******

Royal University of Phnom Penh
{sok.dina, srun.sovila, heng.por, rim.beanbonyka, khoy.saovorak, phaldes}@rupp.edu.kh

Abstract

Text Line Segmentation is one of the important steps for offline printed text or printing recognition and also in constraint Khmer printing text documents remains a challenge because Khmer printing text lines are combined with many consonants, subscript consonants, vowel diacritics and signs and the gap among lines changes following combined words. Each word does not have space separately. In this paper, we proposed a novel technique to segment printing text document into text lines based on the topological assumption that for each text line, there exist a path or the white gaps form left to right that separates those text lines. The purpose of the novel technique has four stages: gray-scaling the input text image, smoothing and noise reducing, components identifying and components merging. To evaluate performance and accuracy of the algorithms mentioned earlier we develop an application called Line Segmentation Reorganization for Khmer printing text by using java platform. Finally, we conduct an experiment with 21 pages that have 400 lines. The result shows that system achieves 95%. We also conduct an experiment with Khmer handwritten text documents with ten native-people with three different categories: four bachelor degree students, three master degree students and three office staffs. The result shows that the system achieves between 93% and 99% for respectively categories.

Keywords: Khmer Texts, Luminosity method, Finite Impulse Response method, Flood Fill Algorithm and Merge bounding box.

1 Introduction

Printing recognition is the transcription of printed data into digital text format. It is traditionally divided into two categories: online recognition and offline recognition. The main difference between them is that the order of the strokes made by writer is available to the online recognition system whereas only the scanned version of the complete printed document is available to offline script recognition. Recently Line Segmentation for printing text has drawn attention to many researchers. Some well-known algorithms such as top-down projection based approach (U. Pal and S. Datta, 2003), smearing approach (S. Messelodi and C. Modena, 1999) and bottom-up component grouping based approach (J. Kumar, L. Kang, D.Doermann, and W. Abd-Almageed, 20011) still remain in challenge for Khmer Language.

In top-down projection based approach (U. Pal and S. Datta, 2003), original image is obtained and then from it the vertical gap between two lines is determined. Then the line-separator is set at the center of the vertical gap. This approach works successfully if the gaps between two lines are significant and the lines are straight enough to resemble machine printed documents. For skewed line segmentation (S. Messelodi and C. Modena, 1999), consecutive black pixels along horizontal direction are smeared. This method requires an appropriate threshold inter each line. The bottom-up component grouping based approach (J. Kumar, L. Kang, D.Doermann, and W. Abd-Almageed, 20011), cannot used for poorly structured documents.

Thus all of the approaches above can't apply with Khmer language, in Khmer language one word is combined with consonants, subscript consonants and vowel diacritics and between consonants and subscript consonants or vowel diacritics don't have space to separate it.

1.1 Problem

Text line segmentation of printing text document is one of the difficult problems in developing a reliable OCR system. Moreover, unlike other language, Khmer language presents unique

challenges such as properties of the line including: line width, height, and space between consonants, subscript consonants or vowel diacritics etc.

1.2 Objectives

The research aims for proposing the new novel method of Text Line Segmentation of Khmer printing text documents are: review and analysis prior proposed method of Text Line Segmentation, Propose a novel method for Text Line Segmentation of Khmer printing text document, develop application by using java platform for demonstration of proposed method, and constitute concept for other researchers or students who want to develop Text Line Segmentation for Khmer language or real Optical Character Reorganization system for Khmer printing text.

1.3 Paper Organization

The remaining parts of this research are organized as follow. Section 2 describe about literature review of other authors such as A Model Based Text Line segmentation method… Section 3 describes the research methodology. This Section separates into five parts. First overview of research methodology is covered. Second describes about the luminosity method for convert RGB image to Grayscale image. Third describes about Loss Pass Finite Impulse Response method for filter dust and make image smooth. Forth, describe about Flood Fill algorithm for segmentation image to text lines. Fifth, describe about merge text line algorithm. Section 4 talks about the experiment, Research Results. Finally, this thesis comes with the conclusion and indicates the future works in section 5.

2 Related Work

2.1 Segmentation of Printed Devnagari Documents

Mr. Vikas J Dongre and Vijay H Mankar at Department of Electronics and Telecommunication, Government Polytechnic, Nagpur, India proposed text-line segmentation algorithm has two main steps (Vikas J Dongre and Vijay H Mankar, 2010). In the first step (Image Preprocessing), they have had four stages. In the first stage (Thresholding and Binarization), they have used histogram based properties of binaries digitized text images into binary images by thresholding using Otsu's method. Original image contains 0 for Object and 1 for background. The image is inverted to obtain image such that object pixels are represented by 1 and background pixels by 0. In the second stage Noise reduction, they are using various morphological processing techniques to eliminate local variations, dilation, erosion and bumps introduced by the optical scanning device or the writing instrument. In the third stage Skew Detection and Correction, Skewed lines are made horizontal by calculating skew angle and making proper correction in the raw image using Hu moments and various transforms. In the fourth stage (Thinning), the boundary detection of image is done to enable easier subsequent detection of pertinent features and objects of interest by using various standard functions. In the second step (Line Segmentation), they have used the global horizontal projection method that has six stages. In the first stage, they construct the horizontal histogram for the image. In the second stage, they count the white pixel in each row. In the third stage, using the Histogram finds the rows containing no white pixel. In the fourth stage, they replace all such rows by 1. In the fifth stage, they invert the image to make empty row as 0 and text lines will have original pixels. In the sixth stage, they mark the bounding box for text lines using standard Matlab functions. However their method can't apply in Khmer Printing Text Line Segmentation. In Khmer language, one word can be combined with consonants, subscript consonants and vowel diacritics. The combination between consonants and subscript consonants is separated in different line as well as the combination between consonants and vowels is also separated. Their method can apply in Khmer documents if they append one method to combine consonants, subscript consonants, vowel diacritics and other signs that are separated in different line to new line.

2.2 A Model Based Text Line Segmentation for Off-Line Handwritten

Mr. Jija Das Gupta at Department of Electronics and Communication Engineering Institute of Technology and Marine Engineering Kolkata, India proposed text-line segmentation algorithm has two main steps (Jija Das Gupta and Bhaba-

tosh Chanda, 2010). In the first step, his system takes the raw image of the given text as an input. After accepting the raw image, all the connected components of image are labeled or identified by using two-pass algorithm. In case there are two characters– one character from a text line and other character from adjacent text line touch each other, so a bounding box large height is formed. To avoid this problem, he set a threshold height (Ht) using the equation, $Ht = \mu + 3\sigma$, where μ is the mean and σ is the standard deviation of the height of existing boundary boxes respectively. Now the bounding boxes with height greater than Ht, are broken into two boxes having half of its previous height. In the second step, the Centroids of all boundary boxes are determined. Centroids of all the components of a line can be arranged properly by using some simple rules based on the following:

1- Two successive bounding boxes must have sufficient vertical overlap
2- The difference in vertical coordinates between two successive centroids is small

Here he defines the overlapping area as

$$\text{Overlapping Area} = \frac{\text{Length h } (h1 \cap h2)}{\text{Min } \{h1, h2\}}$$

Where h1 and h2 are the height of two successive boundary boxes respectively after he found overlapping area, his system arranges the centroids indicated by red dots, so that the text lines are formed. For his two-pass algorithm can find connected component only in the same line or adjacent, with different line or discrete this algorithm can't find currently. His two-pass algorithm must misunderstand when one word combine, one consonant has a lot of subscript consonants and vowel diacritics, it must consider this word combine with two words in different line and it will use equation Ht to separate this word to two words with different line.

3 Research Methodology

In this paper, a new strategy for Khmer printed text line segmentation is proposed. Idea of a text image can be shred into strips along the white gaps in between text lines is motivating. Our approach is based on the topological assumption that for each text line, there exists white gaps or a path from left to right that separate each text line. This method tries to detect such white gaps or paths and use them to shred the image into strips that contain one text line each and this assumption applies to all images containing text in one column layout. Also the white gaps or paths between consonants subscript consonants and vowel diacritics are less than seven pixels and more than seven pixels for white gaps or paths between each text line. The research methodology consists of four main stages namely. In the first stage, the system takes the raw image of the text as an input and converting it to gray-scale image. In the second stage, after accepting the raw image and convert image to grayscale image, grayscale image is filtered by using Loss Pass Finite Impulse Response for making gray-scale image is smooth, reduces noise and reduces high frequency. In the third stage, after grayscale image filtered, all the connected components of image are identified by using Flood Fill algorithm. Then for each connected component a bounding box is found. In the fourth stage, this algorithm merges each bounding box that less than seven pixels into one bounding as can be seen in Figure 3.1.

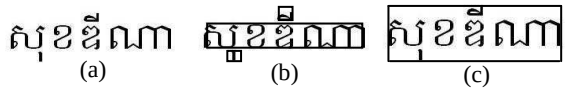


Figure 3.1 (a) Grayscale image of Khmer printed text (b) the bounding box is identified (c) the bounding box less than 7 pixels were merged.

3.1 Convert Original Image to Grayscale

After system accepts original image or RGB image, system converts this original image to grayscale or black and white image by using Luminosity method (Tim Cooper and IrHam , 2011), (John D. Cook, 2009). The luminosity method is a more sophisticated version of the average method. It also averages the values, but it forms a weighted average to account for human perception. We are more sensitive to green than other colors, so green is weighted most heavily. The formula for luminosity is *0.298 R + 0.586 G + 0.114*. This is the sample code for convert RGB image to grayscale image:

```
int r  = ( pixel >> 16)  & 0xff
int g = ( pixel  >> 8 )   & 0xff
int b =  ( pixel & 0xff)
Grayscale = (( r*306)+g* 601)+( b * 117  ))  >> 10
```

Figure 3.1.1 Original Image converted to Grayscale Image (John D. Cook, 2009)

The pixel is combined with three colors Red, Green and Blue. Each pixel is stored as an integer that has four bytes or 32 bits. The first byte is transparency, the second byte is red, the third byte is green and the last byte is blue component. To get red color, it shifts 16 bits of each pixel to the right, which gives now the red is the right most spot, but we also get transparency byte too. To remove transparency byte, we use AND operator with 0xff. Finally we get red color, the same method is used to get green and blue color. After getting red, green and blue color of each pixel, luminosity method is applied to convert each pixel to grayscale pixel by red color multiply with 0.298, green color multiply with 0.586 and blue color multiply with 0.114. It means I get red color 29.8%, Green color 58.6% and Blue color 11.4% of each pixel to convert to grayscale pixel. Finally all colors are summed to get grayscale pixels.

3.2 Filtering Grayscale Image

In this section, grayscale image is filtered by using Loss Pass Finite Impulse Response method (Mitra Basu and Min Su, 2000), (Pankaj Kumar Sa, 2006), (Fundamental of Image processing),

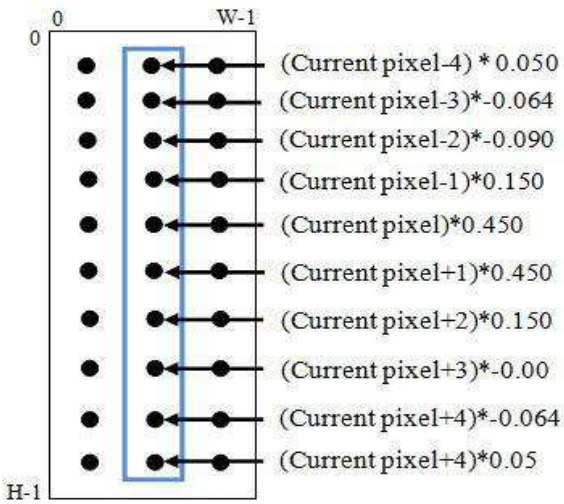


Figure 3.2.1 Filter grayscale Image using vertical FIR

(Ian T. Young, Jan J. Gerbrands and Lucas J. van Vliet, 1998) Loss Pass Finite Impulse Response method reduces luminance variations of luminosity method that is used to convert original image to grayscale image. This method smoothie the image's content, reduces noise effects and removes the spatial high frequency content of the image (details of the image). In this method, it has two main processes: Filter as Horizontal and Filter as Vertical.

3.2.1 Filter as Vertical

Loss Pass Finite Impulse Response method filter image as vertical based on technique using its vertical neighboring area as can be seen in Figure 3.2.1. Let us consider above Figure 3.2.1 of this method. Typically, we get current pixel multiply with 0.450. It means it get current pixel 45%, (current pixel + 1) is 45%, (current pixel+2) is 15%, (current pixel+3) minus 9%, (current pixel+4) is minus 6.4%, (current pixel+4) is 5%, (current pixel-1) is 15%, (current pixel-2) is minus 9%, (current pixel+3) is minus 6.4% and (current pixel-4) is 5%. Finally we sum all to one value of current pixel.

3.2.2 Filter as Horizontal

Loss Pass Finite Impulse Response method filter image as horizontal based on technique using its horizontal neighboring area as can be seen in Figure. 3.2.2. Let us consider above Figure 3.2.2 of this method. Typically, we get current pixel multiply with 0.450. It means it get current pixel 45%,(current pixel + 1) is 45%,(current pixel+2) is 15%,(current pixel+3) minus 9%,(current pixel+4) is minus 6.4%,(current pixel+4) is 5%,(current pixel-1) is 15%,(current pixel-2) is minus 9%,(current pixel+3) is minus 6.4% and (current pixel-4) is 5%. Finally we sum all to one value of current pixel. Here is the formula to compute the Loss Pass Finite Impulse Response as horizontal.

3.3 Text Line Segmentation Algorithm

In this section, I segment the images into the connected components that corresponding to text lines by using Flood Fill algorithm (Lode Vandevenne, 2004), (Soon Tee Teoh, 2004, Lecture 4). Flood Fill algorithm, also called seed fill, is an algorithm that determines the area

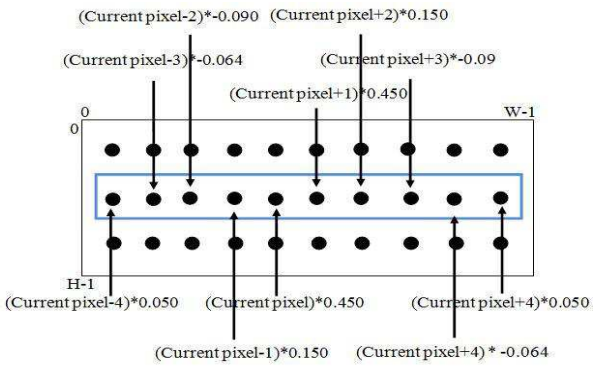


Figure 3.2.2 Filter grayscale image using Horizontal FIR

connected. When Flood Fill algorithm is applied on an image to fill a particular bounded area with color, it is also known as boundary fill. From concept of this algorithm, we assume the Grayscale image is stored in array I(x,y) which contains single column text information:

$$I(x,y) = \begin{cases} \text{Black/Foreground pixel if each pixel} < \text{whiteThreshold} \\ \text{Whitespace/background pixel if each pixel} >= \text{whiteThreshold} \end{cases}$$

Where x € [1, I_{width}] and y € [1, I_{height}] and whiteThreshold is 128 is the mean or media of grayscale color. This whiteThreshold uses to determine black/foreground pixel and whitespace/background pixel. We also assume that any pair of consecutive text lines, there exist a path/white gap from left to right that separates those two text lines. This algorithm tries to detach such path or white gaps and use them to segment the image to bounding box, here is the algorithm to segment the grayscale image to bounding boxes as can be seen in Figure 3.3.1. Let us consider above Figure 3.3.1 of the algorithm. Typically, the algorithm is working from y=0 axes to height-1(top to button). For each y axes, it will loop form x=0 axes to width -1 (left to right) to find black pixels and white space. The algorithm has two main stages. In the first stage, it tries to detect black pixel and stores it in array one dimension. This black pixel is represented as y1 or starting point of connected component of each text lines. In the second stage, it will skip all black pixels and try to detect whitespace pixel. After detecting whitespace pixel, it stores whitespace pixel in array one dimension. Whitespace pixel is represented as y or end point of connected component of the text lines as can be seen in bellow Figure 3.3.2 and Figure 3.3.3.

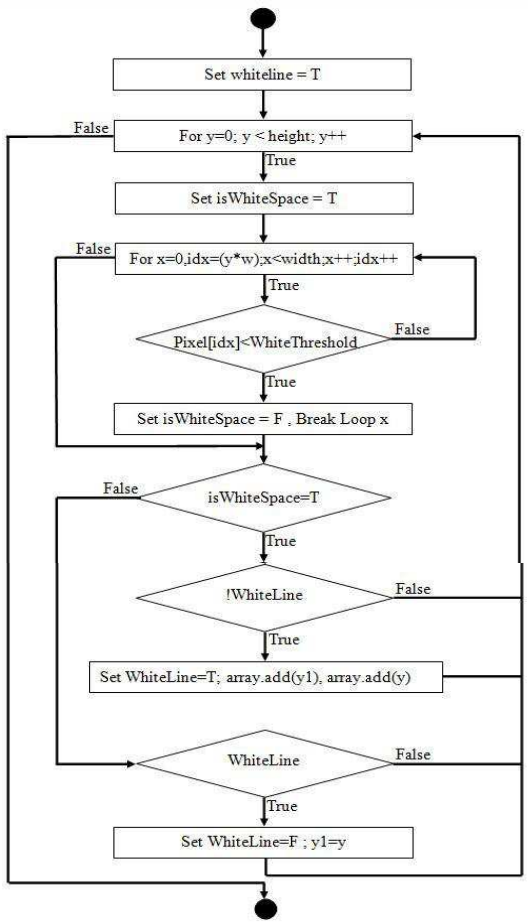


Figure 3.3.1 Text Line Segmentation Algorithm

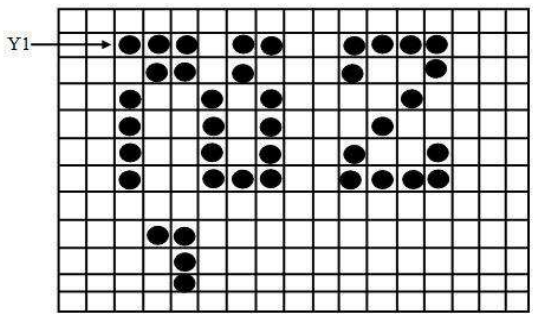


Figure 3.3.2 Stage 1, black pixel is detected and store in y1

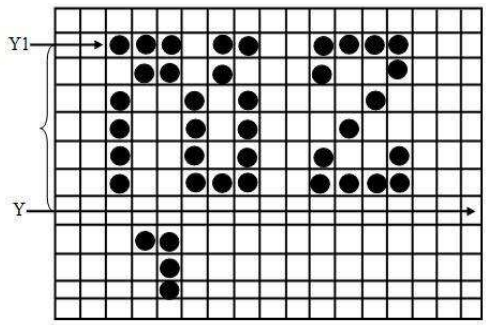


Figure 3.3.3 stage 2, Whitespace is detected and store in y. From y1 to y is represented one connected component

3.4 Merge Bounding Boxes to New boxes

In this section, we merge bounding boxes that got from Flood Fill algorithm section 3.3 to the new bounding box. For this algorithm, Whitegap between two bounding boxes that less than 7 pixels; it will be merged to the new bounding box. Here is the algorithm to merge the bounding boxes to the new bounding box can be seen in Figure 3.4.1.

Let consider Figure 3.4.1 of the algorithm. Typically, the algorithm gets two bounding boxes to compare. If Whitegap between the first bounding box and the second bounding box less than 6% of the high of the first bounding box or Whitegap between the first bounding box and the second bounding bound less than seven pixels, it will be considered as consonants, subscript consonants or vowel diacritics and it will be merge to the new bounding box or text line which can be seen in Figure 3.4.2.

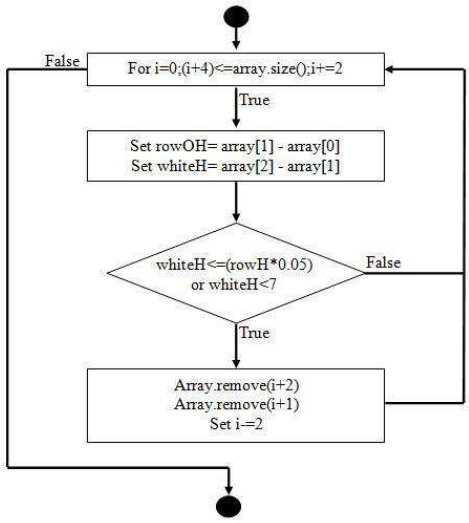


Figure 3.4.1 Merge bounding box Algorithm

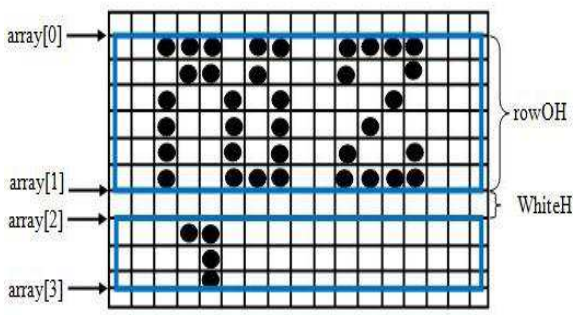


Figure 3.4.2 from array [0] to array [1] is represented the first bounding box, from array [2] to array [3] is represented the second bounding box

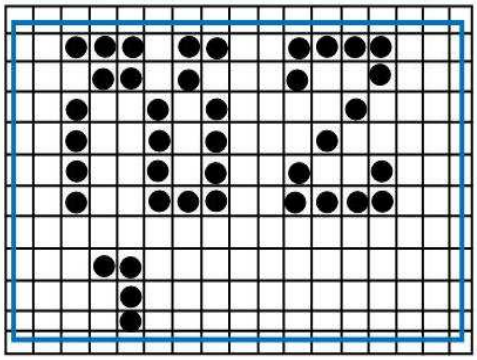


Figure 3.4.3 the first bounding boxes and the second bounding box are merged to the new one bounding box

Let consider Figure 3.4.2. If WhiteH between the first bounding box and the second bounding box less than 6% of height of the first bounding box or WhiteH between the first bounding box and the second bounding box less than seven pixels. The first bounding box and the second bounding box are merged to the new one bounding box as can be seen in Figure 3.4.3.

In this paper, we present four main algorithms for recognizing Line Segmentation for Khmer printing text documents, English printing text documents and Khmer handwriting text documents which are Luminosity Algorithm method (Tim Cooper and IrHam, 2011), (John D. Cook, 2009) for convert original image to Grayscale image, Loss Pass Finite Impulse Response algorithm (Mitra Basu and Min Su, 2000), (Pankaj Kumar Sa, 2006), (Fundamental of Image processing), (Ian T. Young, Jan J. Gerbrands and Lucas J. van Vliet, 1998) for reducing luminance variations of luminosity algorithm and making Grayscale image smooth (FIR algorithm has two main processes: Filter as Horizontal and Filter as Vertical), Flood Fill Algorithm (Lode Vandevenne, 2004), (Soon Tee Teoh, 2004, Lecture 4) for segmenting the grayscale images into the connected components that corresponding to text lines and Merge Bounding boxes algorithm for merge two lines to one new line when Whitegap between the first bounding box and the second bounding box less than 6% of height of the first bounding box or Whitegap between the first bounding box and the second bounding box less than seven pixels. We conclude that these four algorithms are applicable for Khmer and English printing documents with the difference satisfaction based on level of the accuracy.

4 Research Results

4.1 Experiment

In order to evaluate the accuracy and performance of Line Segmentation for Khmer printing text documents, Line Segmentation for Khmer printing text Application by using java platform is developed. Testing is conducted on English and Khmer printing text documents on 25 pages equal 400 lines. We also test on Khmer handwriting text document with ten people (4 bachelor degree, 3 master degree and 3 office staffs).

4.2 Results and Evaluation

For the experiment Khmer printing documents and English printing documents, the total number of lines is 400 lines of 21 pages. For the experiment Khmer handwriting line segmentation documents, each participant is asked to write 9 lines as Khmer language. Therefore, the total number of lines of each participant is 9 lines multiplied by 10 participants thus we got 90 tested lines. We evaluate Khmer line segmentation documents by Accuracy. The level of recognition accuracy is calculated by:

$$\text{Accuracy} = \frac{\text{Total of Correct Lines}}{\text{Total of Lines}}$$

$$\text{Result} = \frac{\text{380 of correct lines}}{\text{400 of total lines}} = 95\ \%$$

Table 4.2.1 Show accuracy of Khmer handwriting line segmentation from each user:

Participants	Accuracy
User 1	99%
User 2	95%
User 3	96%
User 4	99%
User 5	97%
User 6	96%
User 7	98%
User 8	95%
User 9	95%
User 10	96%

Table 4.2.1 Accuracy of Khmer handwriting

5 Conclusion

In this paper, we present one novel technique for Text line segmentation of Khmer printing documents that has four main algorithms: luminosity algorithm for convert original image to grayscale image, Loss Pass Finite Impulse Response algorithm for filter noise, for making image smooth, for reducing high frequency, Flood Fill algorithm for line segmentation and Merge line algorithm. We conduct a test with 21 pages (400 lines) for Khmer and English printing documents and 10 people with total line 9 for each person of Khmer printing documents. The overall result shows that 95% for Khmer and English printing documents and 93-99% for Khmer handwriting documents achieved by this method. From this result, we conclude that this method is applicable for Line Segmentation of Khmer language based on the accuracy level of testing.

In our future work, we would like develop system that will be able to recognize for both Khmer word and character with high accuracy result. For Khmer handwriting, Khmer word or character touching, overlapping components from one line to other line should be considered.

References

[1] A.Nicolaou,B. Gatos. (2009). *Handwritten Text Line Segmentation by Shredding Text into its Lines.* International Conference on Document Analysis and Recognition.

[2] C. Weliwitage, A. L. Harvey and A. B. Jennings. (2005). *Handwritten Document Offline Text Line Segmentation.*

[3] Cote, Guy (San Jose, CA, US) Frederiksen and Jeffrey E. (Sunnyvale, CA, US). (2011). *System and Method for Deteching and Correcting Defective Pixels in an image sensor.*

[4] Ian T. Young, Jan J. Gerbrands and Lucas J. van Vliet. (1998). *Fundamental of Image processing.* Delft university of Technology.

[5] Jija Das Gupta and Bhabatosh Chanda. (2010). *A Model Based Text Line Segmentation Method for Off-Line Handwritten Documents.* International Conference on Frontiers.

[6] Mitra Basu and Min Su (2000). *Image Smoothing with Exponential Functions* .

Engineering Department, the City College of the City University of New York, New York, N. Y. 10031

[7] Naresh Kumar Garg, Lakhwinder Kaur and M. K. Jindal. (2010). *A New Method for Line Segmentation of handwritten Hindi Text* . International Conference.

[8] Pankaj Kumar Sa. (2006). *On the Development of Impulsive noise Removal Schemes*. Department of Computer Science & Engineering Institute of Technology, India.

[9] Soon Tee Teoh. (2007), Lecture 4: *Flood Fil Algorithm.* CS 11B [PowerPoint Slides]

[10] Vikas J Dongre and Vijay H Mankar (2011). *Segmentation of Printed Devnagari Documents Mitra Basu and Minsu.*

[11] Yangdong Gao, Xiaoqing Ding and Changsong Liu (2011). *A Multi-scale Text Line Segmentation Method in Freestyle*. International Coference on Document Analysis and Recognition.

[12] Jendem (2013). *Finite Impulse Response Method*. Retrieved from http://en.wikipedia.org/wiki/Finite_impulse_response

[13] Z. Lu, R. Schwartz, P. Natarajan, I. Bazzi, and J. Makhoul, " Advances in the bbn byblos ocr system," in ICDAR, sept. 1999

[14] R. Haralick, "Document image understanding: geometric and logical layout," in CVPR, June 1994

[15] S. Messelodi and C. Modena, "Automatic identification and skew estimation of text lines in real scene images,"Patter Recognition, vol. 32, no. 5, 1999

[16] J. Kumar, L.Kang, D. Doermann, and W. Abd-Almageed, " Segmentation of handwritten textlines in presence of touching components" in ICDAR, sept. 2011, pp. 109-113.

[17] U. Pal and S. Datta, "Segmentation of bangla unconstrained handwritten text," in ICDAR, Aug. 2003, pp. 1128-1132

A Comparative Study on Confidence-Weighted Learning between Fuzzy and Non-Fuzzy Classifiers

Takeshi Sumitani† **Tomoharu Nakashima**† **Andrzej Bargiela** ‡
†Graduate School of Engineering
Osaka Prefecture University, Japan
‡School of Computer Science
The University of Nottingham, U.K.
takeshi.sumitani@ci.cs.osakafu-u.ac.jp
tomoharu.nakashima@kis.osakafu-u.ac.jp
andrzej.bargiela@nottingham.ac.uk

Abstract

The performance of fuzzy and non-fuzzy on-line classifiers are examined in this paper. Confidence-weighted learning is employed as an on-line algorithm. In the case of non-fuzzy classifiers, the linear combination of input attributes is used to determine the class of the input pattern. The weights in the linear combination are modelled using Gaussian distributions. The mean and variance values are modified so that the new training patterns are correctly classified while remembering the information of the previous training patterns. On the other hand, a fuzzy classifier comprises a set of fuzzy if-then rules, in each of which the membership value for the correspondent attribute value is calculated. The membership values for an input vector with the fuzzy if-then rules are then used as an input vector in the linear combination. The weight in the linear combination is specified in the consequent part of each fuzzy if-then rule. A series of computational experiments are conducted to examine the classification performance of the fuzzy and the non-fuzzy on-line learning methods for large scale real-world classification problems that are available in the UCI repository.

Keywords: Fuzzy if-then rules, On-line learning, Pattern classification, Linearly weighted function

1 Introduction

On-line learning plays an important role specifically in a dynamic problem where the decision making for a situation changes over time. It is also useful when a problem size is too big to handle at one time. In this case, it is necessary for a decision making model to learn incrementally using a stream of training patterns where only a portion of the training data set is available at each time step.

There are several approaches to the on-line learning. For example, Crammer et al.[1] proposed passive-aggressive algorithm where parameters of a classification model is updated after a new set of training patterns are available so that the new patterns are correctly classified with a specified maximum error. For fuzzy rule-based classifiers, Nakashima et al.[5; 6; 7] proposed incremental learning algorithms where the internal parameters for determining the weight of the fuzzy if-then rules are updated according to the new set of training patterns.

Dredze et al.[4] proposed a confidence-weighted learning algorithm for linear classifiers. In the confidence-weighted learning algorithm, the weights in the linear discriminant function are modelled not by a real value but by a Gaussian function. The parameters of the Gaussian function such as a mean and a variance is updated so that the amount of the update in Gaussian functions is minimum in terms of KL-divergence. The variance of a Gaussian function represents the confidence of the weight. The larger the variance is, the less confident the weight is. Those weights whose variance of the Gaussian function is large are largely updated while those with a small variance are updated only a little as the weights are already confident.

Fuzzy rule-based systems involves a set of fuzzy if-then rules that represents the domain knowledge. Fuzzy if-then rules can be generated manually through the interview with the domain experts although it is sometimes difficult for the domain experts to quantitatively and qualitatively explain their expertise. If there are

available numerical data that represent the domain knowledge, there is a room for learning to automatically generate fuzzy if-then rules from them. The automatic generation of fuzzy if-then rules has been investigated in soft computing community for the last 20 to 30 years. In this paper, an automatic generation method of fuzzy if-then rules is considered to construct fuzzy classifiers.

Fuzzy classifiers have been known for their high classification performance for high-dimensional non-linear problems [2]. A fuzzy classifier is composed of a set of fuzzy if-then rules, which has its antecedent part and consequent part. The antecedent part of fuzzy if-then rules represents a fuzzy covered area in the input space. Each fuzzy set in the antecedent part has a linguistic label (such as "big" and "small") that shows a fuzzy supported interval in the correspondent attribute axis. Therefore, it is possible to linguistically understand the mapping that is achieved by the fuzzy classifier. This is one of the characteristic feature of fuzzy classifiers. This characteristic feature allows to acquire linguistic knowledge from numerical data. The linguistic knowledge acquisition using fuzzy classifiers is one of active research areas in data mining and soft computing [3]. However, there are not many articles that deal with the combination of fuzzy classifiers and on-line learning, especially the confidence-weighted learning.

In this paper, the classification performance of fuzzy and non-fuzzy on-line classifiers are examined. For the on-line learning algorithm, the confidence-weighted learning method is used for both fuzzy and non-fuzzy classifiers. A series of computational experiments are conducted using several large-scale real-world classification problems that are available from the UCI repository [10].

2 Confidence-Weighted Learning for Linear Discriminant Functions

This section describes a confidence-weighted learning algorithm for linear discrimination. First the standard procedure of classification using a linear discrimination function is explained, which is followed by the explanation of the confidence-weighted learning algorithm for training the linear discriminant function. In this paper only binary classification is considered where the number of class labels is just two. An extension to the classifiers for allowing multi-class classification is not our focus of this paper, but will be addressed in our future work.

2.1 Linear Discriminant Function

Let us define without losing generality that an input pattern is represented as an n-dimensional real vector in a hyper cube $[0.0, 1.0]^n$. It is also assumed that the n-dimensional weight vector $\vec{w} = (w_0, w_1, \ldots, w_n)$ is already specified. An unseen pattern $\vec{x} = (x_1, x_2, \ldots, x_n)$ is classified as follows:

$$\begin{cases} \vec{x} \text{ is Class 1}, & \text{if } y > 0.0, \\ \vec{x} \text{ is Class 2}, & \text{if } y < 0.0, \\ \text{Classification rejected}, & \text{otherwise}, \end{cases} \quad (1)$$

where

$$y = w_0 + w_1 \cdot x_1 + \ldots + w_n \cdot x_n. \quad (2)$$

Thus, the classification of unseen patterns is determined according to the sign of the inner product of the weight vector and the augmented input vector $\vec{x'} = (1, x_1, x_2, \ldots, x_n)$.

2.2 Confidence-Weighted Learning

In the confidence-weighted learning, each element (i.e., each weight) in the weight vector $\vec{w} = (w_0, w_1, \ldots, w_n)$ is modelled by a Gaussian function with a mean and a variance. Let us denote that the mean value of the i-th Gaussian function for modelling w_i is μ_i and its variance is σ_i. Then the weights become random variables of the probability distribution with specified mean and variance. When a set of new training patterns become available, the Gaussian distributions are updated so that the classification of the new training patterns are more reliable with a minimum amount of update. During the course of the algorithm execution, the variances of the Gaussian distributions gradually become small.

The update formulation of the Gaussian functions for the weight vector is obtained by solving the following optimization problem:

$$(\overrightarrow{\mu^w_{\text{new}}}, \overrightarrow{\Sigma^w_{\text{new}}}) = \underset{(\overrightarrow{\mu_t}, \overrightarrow{\Sigma_t})}{\arg\min} D_{KL} \left(N(\overrightarrow{\mu_t}, \overrightarrow{\Sigma_t}) \parallel N(\overrightarrow{\mu^w_{\text{old}}}, \overrightarrow{\Sigma^w_{\text{old}}}) \right), \quad (3)$$

subject to:

$$Pr_{\overrightarrow{w} \sim N(\overrightarrow{\mu}, \overrightarrow{\Sigma})} \left[y_t(\overrightarrow{w} \cdot \overrightarrow{x}) \geq 0 \right] \geq \eta, \quad (0.5 \leq \eta \leq 1.0), \tag{4}$$

where y_t represents the target output value (1 or -1). $\overrightarrow{\mu^w}$ and $\overrightarrow{\Sigma^w}$ are the mean and the variance vectors for modelling $\vec{w}$, respectively. From this optimization problem, the following equation is obtained:

$$\overrightarrow{\mu^w_{\text{new}}} = \overrightarrow{\mu^w} + \alpha \cdot y_t \cdot \overrightarrow{\Sigma^w} \cdot \overrightarrow{x}, \tag{5}$$

$$\overrightarrow{\Sigma^w_{\text{new}}}^{-1} = \overrightarrow{\Sigma^w}^{-1} + 2 \cdot \alpha \cdot y_t \cdot \phi \cdot \text{diag}(\overrightarrow{x}), \tag{6}$$

$$\alpha = \max\{\gamma, 0\}, \tag{7}$$

$$\gamma = \frac{-(1+2\phi M)}{4\phi V} + \frac{\sqrt{(1+2\phi M)^2 - 8\phi(M-\phi V)}}{4\phi V}, \tag{8}$$

$$M = y_t \cdot (\overrightarrow{x} \cdot \overrightarrow{\mu^w}), \tag{9}$$

$$V = \overrightarrow{x}^\top \cdot \overrightarrow{\Sigma^w} \cdot \overrightarrow{x}, \tag{10}$$

$$\phi = \Phi^{-1}(\eta), \tag{11}$$

where $\Phi()$ represents the cumulative distribution function of the standard normal distribution. The common specification of the value ϕ is 1.0 in the literature instead of computing it from the probability η. In this case, η is calculated as $\eta \approx 0.84$.

The update procedure is shown in the following:

Step 1: Initialize $\overrightarrow{\mu^w}$ and $\overrightarrow{\sigma^w}$.

Step 2: Obtain a new training pattern $\vec{x}$.

Step 3: Update the mean and variance using Equations (18) and (19).

Step 4: Classify unknown patterns using $\overrightarrow{y}$ to evaluate the performance of the current fuzzy classifier. If any termination criterion are not satisfied, return to Step 3. Otherwise, halt the procedure.

3 Confidence-Weighed Learning for Fuzzy Classifiers

The original work of the on-line learning was applied to linear classification where the attribute values of an input value are linearly combined using Gaussian modelled weights. When a set of new training patterns are obtained, the correspondent Gaussian distributions to the weights are updated so that the linear classifier correctly classify not only the new patterns but also the previous training patterns with least amount of modification. During the course of the algorithm execution, the variances of the Gaussian distributions gradually become small as the confidence of the weights gets higher.

This section describes a fuzzy classifier and its on-line adaptive procedure using confidence-weighted learning. A fuzzy classifier consists of a set of fuzzy if-then rules. As in the last section, it is assumed that the pattern classification problem has the dimensionality n and two classes.

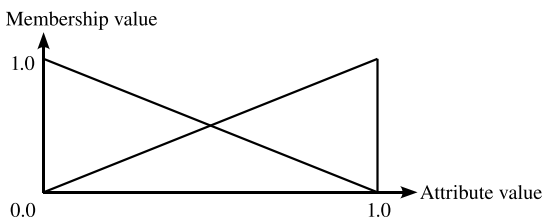
3.1 Fuzzy If-Then Rules

The following type of fuzzy if-then rules is used for our fuzzy classifier:

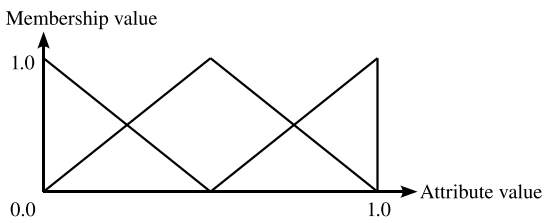
$$R_q: \text{ If } x_1 \text{ is } A_{q1} \text{ and } \cdots \text{ and } x_n \text{ is } A_{qn} \text{ then } y_q, \quad q = 1, 2, \cdots, N \tag{12}$$

where R_q is the label of the q-th fuzzy if-then rule, $\overrightarrow{A_q} = (A_{q1}, A_{q2}, \cdots, A_{qn})$ represents a vector of antecedent fuzzy sets, y_q is the consequent real value and N is the total number of generated fuzzy if-then rules. Triangular membership functions are used for antecedent fuzzy sets in this paper. The number of triangular fuzzy membership functions for each attribute must be specified before a fuzzy classifier is constructed. Examples of the triangular fuzzy partition are shown in Figure 1. In Figure 1, two examples of fuzzy partitions are shown where the attribute is divided into two and three triangular fuzzy sets. It should be noted that the fuzzy partition has the effect on the complexity of constructed fuzzy classifiers. The number of generated fuzzy if-then rules in constructing a fuzzy classifier depends on the fuzzy partition for attributes. For example, if the number of triangular fuzzy membership functions is two for a ten-dimensional problem, the number of fuzzy if-then rules is $2^{10} = 1024$ while it is $3^{10} = 59049$ when the number of triangular

membership functions is three for each attribute. Although a fuzzy classifier become intractably complicated with a huge number of fuzzy if-then rules, it is able to realize highly non-linear mapping with high accuracy. On the other hand, it is difficult to linguistically understand the input-output mapping that the constructed fuzzy classifier achieves. Thus, the number of fuzzy membership functions should be specified carefully considering the balance between the accuracy and the complexity.



(a) Two fuzzy sets.



(b) Three fuzzy sets.

Figure 1. Triangular fuzzy membership functions.

3.2 Classification of Unknown Patterns

The classification of an unknown pattern is done using the output value that are calculated through fuzzy inference process. In this paper, we classify unknown patterns $\overrightarrow{x} = (x_1, x_2, \cdots, x_n)$ as follows:

$$\begin{cases} \vec{x} \text{ is Class 1}, & \text{if } y > 0.0, \\ \vec{x} \text{ is Class 2}, & \text{if } y < 0.0, \\ \text{Classification rejected}, & \text{otherwise}, \end{cases} \quad (13)$$

where

$$y = y_1 \cdot \mu_1 + y_2 \cdot \mu_2 + \ldots + y_N \cdot \mu_N, \quad (14)$$

where y_q is the consequent real value for the q-th fuzzy if-then rule and $\mu_q(\vec{x})$, $q = 1, 2, \ldots, N$ is the membership value of $\vec{x} = (x_1, x_2, \ldots, x_n)$ with the q-th fuzzy if-then rule R_q. It is calculated as follows:

$$\mu_q(\vec{x}) = \mu_{A_{q1}}(x_1) \cdot \mu_{A_{q2}}(x_2) \cdot \ldots \cdot \mu_{A_{qn}}(x_n), \quad (15)$$

where $\mu_{A_{qi}}(x_i)$, $i = 1, 2, \ldots, n$, $q = 1, 2, \ldots, N$ is the membership value of x_i for the i-th attribute with the rule R_q.

3.3 Updating Consequent Real Values by CW Learning

The confidence-weighted learning is realized for fuzzy classifiers by taking the consequent real values in (14) as the weight in the linear classification (i.e., $w_0, w_1, \ldots, w_n$ in (2)). Thus, in the case of the fuzzy classifiers, the consequent real values in(12) are modelled using Gaussian distributions so that the value y_q of rule R_q is treated as the mean μ_q of a Gaussian distribution with the variance σ_q. Following the case for the linear classification, the update rule of consequent real values is obtained by solving the following optimization problem::

$$(\overrightarrow{\mu^y_{\text{new}}}, \overrightarrow{\Sigma^y_{\text{new}}}) = \arg\min_{(\overrightarrow{\mu_t}, \overrightarrow{\Sigma_t})} D_{KL}\left(N(\overrightarrow{\mu_t}, \overrightarrow{\Sigma_t}) \parallel N(\overrightarrow{\mu^y_{\text{old}}}, \overrightarrow{\Sigma^y_{\text{old}}})\right) \quad (16)$$

subject to

$$Pr_{\overrightarrow{y} \sim N(\overrightarrow{\mu}, \overrightarrow{\Sigma})}\left[y_t(\overrightarrow{y} \cdot \overrightarrow{x}) \geq 0\right] \geq \eta, \quad (17)$$
$$(0.5 \leq \eta \leq 1.0)$$

where y_t represents the target output value (1 or -1). $\overrightarrow{\mu^y}$ and $\overrightarrow{\Sigma^y}$ are the mean and the variance of $\overrightarrow{y}$, respectively. From this optimization problem, the following equations are obtained:

$$\overrightarrow{\mu^y_{\text{new}}} = \overrightarrow{\mu^y_{\text{old}}} + \alpha y_t \overrightarrow{\Sigma^y} \overrightarrow{x}, \quad (18)$$

$$\overrightarrow{\Sigma^y_{\text{new}}}^{-1} = \overrightarrow{\Sigma^y_{\text{old}}}^{-1} + 2\alpha \cdot y_t \cdot \phi \cdot \text{Diag}(\overrightarrow{x}), \quad (19)$$

$$\alpha = \max\{\gamma, 0\}, \quad (20)$$

$$\gamma = \frac{-(1 + 2\phi M)}{4\phi V} + \frac{\sqrt{(1 + 2\phi M)^2 - 8\phi(M - \phi V)}}{4\phi V}, \quad (21)$$

(22)

Table 1. Details of the data sets that are used in the computational experiments.

Data set	# of patterns	Dimensionality	Classes
Abalone	4177	8	2
Breastcancer	609	10	2
Pima-indians-diabetes	768	8	2
Transfusion	748	4	2

$$M = y_t(\overrightarrow{x} \cdot \overrightarrow{\mu^y}), \quad (23)$$

$$V = \overrightarrow{x}^\top \overrightarrow{\Sigma^y} \overrightarrow{x}, \quad (24)$$

$$\phi = \Phi^{-1}(\eta), \quad (25)$$

where $\Phi()$ represents the cumulative distribution function of the standard normal distribution as in the case of the linear classification. The update procedure is also the same as in the update procedure of the linear classifiers.

4 Computational Experiments

This section examines the performance of the fuzzy and non-fuzzy classifiers with the confidence-weighted learning algorithm. In the experiments of this paper, classifiers are applied to a set of large-scale real-world pattern classification problems that are available from UCI repository [10].

This subsection deals with static but large-scale classification problems. High-dimensional data sets with a large number of training patterns are picked up from the UCI repository [10]. The details of the data sets that are used in the computational experiments in this subsection are shown in Table 1. Since the fuzzy and linear classifiers considered in this paper are designed for two-class pattern classification problem, the selected date sets have only two classes.

4.1 Experimental Settings

Since the purpose of the experiments in this subsection is to examine the classification performance of classifiers with on-line learning, not all training patterns are given for constructing a classifier but only one training pattern becomes available at each time step. Thus, a classifier must be constructed incrementally for each new training pattern. The procedure of this experimental setting is shown below:

Step 1: Randomly specify the parameters of the classifiers (i.e., the weights in the case of the linear classification and the consequent real values in the case of the fuzzy classifiers).

Step 2: Randomly select a training pattern from the training set without replacement.

Step 3: Update classifiers using the selected training pattern.

Step 4: If all the training patterns are selected, terminate the procedure. Otherwise go to Step 2.

The classification performance is evaluated by averaged accuracy. Ten-fold cross-validation was iterated ten times. In the ten-fold cross validation, the entire data set is divided into ten groups and each group is used as test patterns for examining the generalisation ability of the classifiers. This procedure is iterated until all the ten groups are used as test patterns just once. The averaged classification rate for ten trials of the ten-fold cross validation is calculated to use as a performance measure.

4.2 Experimental Results

The results of the experiments are shown in Table 2. Table 2 shows the average classification rates of the fuzzy and linear classifiers on the data sets in Table 1. From the experimental results, it cam be seen that the fuzzy classifier achieved higher classification rate than the linear classifier for three data sets out of the four.

Table 2. Experimental results (Average classification rates over ten trials of 10-CV).

Data set	Fuzzy	Linear
Abalone	0.8772	0.8671
Breastcancer	0.9591	0.9213
Pima-indians-diabetes	0.6766	0.5706
Transfusion	0.7061	0.7319

5 Conclusions

This paper investigates the classification performance of the fuzzy and the non-fuzzy classifiers. Both types of classifiers adopt Confidence-Weighted (CW) learning where the weight parameters in the classifiers are modified so that new training patterns are correctly classified while as many previous patterns as possible are also correctly classified.

The common characteristic feature of the fuzzy and the non-fuzzy classifiers is that both classifiers use a linear combination to obtain the final classification results. In the non-fuzzy classifier, the input values are summed up with their associated weights while in the case of fuzzy classifiers the membership values of fuzzy if-then rules that constitute the classifier are linearly combined. Thus, the number of weight parameters for the fuzzy classifiers are higher than that of non-fuzzy classifiers in general. The advantage of fuzzy classifiers over non-fuzzy ones is that the input-output mapping can be linguistically understood at the cost of a large number of weight parameters.

The computational experiments in this paper used a large-scale pattern classification problems that are available in the UCI repository. It was shown that the generalization ability of the fuzzy classifiers is higher than that of the non-fuzzy classifiers. Thus, it can be said that the fuzzy classifiers have high adaptation ability for incremental classification problems where all training patterns can not be used for constructing a classifier at once.

Future works include other formulation of fuzzy classifiers such as one-class classification and regression problems.

References

[1] K. Crammer and O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, "Online Passive-Aggressive Algorithms," *Journal of Machine Learning Research*, Vol. 7, pp. 551–585, 2006.

[2] H. Ishibuchi, T. Nakashima and M. Nii,*Classification and Modeling with Linguistic Information Granules*, Springer, 2003.

[3] L. I. Kuncheva,"How Good Are Fuzzy If-Then Classifiers," IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics, vol. 30, no. 4, pp. 501–509, August, 2000.

[4] M. Dredze, K. Crammer, and F. Pereira,"Confidence-Weighted Linear Classification,"*Proc. of the 25th International Conference On Machine Learning*, pp. 264–271, 2008.

[5] T. Nakashima, T. Sumitani, and A. Bargiela, "Incremental Update of Fuzzy Rule-Based Classifiers for Dynamic Problems," *Computer and Information Science 2012*, pp.209–219, 2012.

[6] T. Nakashima, T. Sumitani, and A. Bargiela, "Incremental Learning of Fuzzy Rule-Based Classifiers for Large Data Sets," *Proc. of 2012 World Automation Congress*, CD-ROM, 2012.

[7] T. Nakashima, T. Sumitani, and A. Bargiela, "Performance Evaluation of Incremental Fuzzy Rule-Based Classifiers," *Computers, Networks, Systems and Industrial Application,* pp.234–241, 2012.

[8] T. Nakashima, Y. Kuroda, Y. Higuchi, Y. Maekawa, "Occupancy Simulation in an Elderly Monitoring Environment," *Proc. of SIG-ECOmp, SOFT*, pp.30–34, 2013.

[9] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern classification, 2nd edition*, Wiley, New York, Wiley, 2001.

[10] A. Asuncion and D. J. Newman, UCI Machine Learning Repository, `http://www.ics.uci.edu/~mlearn/MLRepository.html`, 2007

Extending *k*-Representative Clustering Algorithm with an Information Theoretic-based Dissimilarity Measure for Categorical Objects

Thu-Hien Thi Nguyen† **Van-Nam Huynh**†
†School of Knowledge Science
Japan Advanced Institute of Science and Technology, Japan
{nguyen.hien,huynh}@jaist.ac.jp

Abstract

This paper aims at introducing a new dissimilarity measure for categorical objects into an extension of k-representative algorithm for clustering categorical data. Basically, the proposed dissimilarity measure is based on an information theoretic definition of similarity introduced by Lin [15] that considers the amount of information of two values in the domain set. In order to demonstrate the efficiency of the extended k-representative algorithm with the new dissimilarity measure, we conduct a series of experiments on real datasets from UCI Machine Learning Repository and compare the result with several previously developed algorithms for clustering categorical data.

Keywords: Cluster Analysis, Categorical Data Clustering, K-means, Dissimilarity measure

1 Introduction

During the last two decades or so, data mining has emerged as a rapidly growing interdisciplinary field, which merges together databases, statistics, machine learning and other related areas in order to extract useful knowledge from data [1]. Cluster analysis or simply clustering is one of fundamental tasks in data mining that aims at grouping a set of data objects into multiple clusters, such that objects within a cluster are similar one another, yet dissimilar to objects in other clusters. Dissimilarities and similarities between objects are assessed based on those attribute values describing the objects and often involve distance measures.

Typically, objects can be considered as vectors in n-dimensional space, where n is the number of features. When objects are described by numerical features, the distance measure based on geometric concept such as Euclid distance or Manhattan distance. But these geometric distance measures are not applicable for categorical data which contains values from gender, locations, etc. Recently, clustering data with categorical attributes have gained increasing attention [9; 7; 10; 8; 4; 5]. As for categorical data, the comparison measure is most naturally used [4]. However, this metric does not distinguish between the different values taken by the attribute, since we only measure the equality between pair of values, as argued in [16].

In this paper we propose a new approach to compute the dissimilarity between values of a categorical variable and apply it to k-representative clustering of categorical data. The solution is to consider the amount of information between values in the domain set of a categorical attribute. The closer the amount of information between values are, the more similarities they share [15].

The remainder of this paper is organized as follows: Section 2 briefly describes related work. Section 3 recalls the notations and definitions used in this paper. Section 4 deals with the function of k-means algorithm. The proposed method is discussed in Section 5. In Section 6, we present the results of the experiments on data sets. Finally, Section 7 concludes the paper.

2 Related work

Probably, the k-means clustering [2] is the most well-known approach for clustering data sets with numerical attributes. It is a traditional partitioning based approach which starts with k random centroids and the centroids are updated iteratively by computing the average of the numerical features in each cluster. Each observation or object is assigned to clusters based upon the nearest distance to the means of the clusters. The iteration continues until the assignment is stable, that is, the clusters formed in the current round are the same as those formed in the previous round. The k-means is very popular due to its ease of implementation, linear time complex-

ity in size of the data, and almost surely convergence to local optima [3]. However, in real life many data set are categorical. In such situations, k-means cannot be applied.

Few algorithms have been proposed in recent years in order to overcome the numerical-only limitation of k-means, like k-modes [5], k-representative [13]. Huang proposed k-modes, an extension of k-means, which is developed to cluster categorical data. Basically, the k-modes algorithm uses a simple matching similarity measure for categorical objects, and modes instead of means for cluster centers. Mode is a data point, in which the value of each attribute is assigned the most frequent value in the attribute's domain set. Further, Huang also combined the k-modes algorithm with k-means algorithm in order to deal with mixed numerical and categorical databases. These extensions allow us to efficiently cluster very large data sets from real world applications. However, the performance of k-mode algorithm depend strongly on the selection of the modes during the clustering process. In an attempt to overcome this drawback, San et al. [13] introduced a new dissimilarity measure to the k-modes clustering process to improve the accuracy that resulted in the so-called k-representative algorithm. The main idea of k-representative algorithm is to use the relative attribute frequencies of the cluster modes in the similarity measure in the k-modes objective function. It has been shown that the k-representative algorithm is very effective [14].

Along with trying to adapt k-means for clustering categorical objects, other approaches such as ROCK [10], COOLCAT [12], LIMBO [11] have been developed. Particularly, ROCK [10] uses the user-defined threshold of Jaccard similarity measure between the two objects to define the links between them. Then it tries to optimize cluster objective function with respect to the number of links between the cluster members. The basic algorithm is hierarchical and not suitable for large datasets due to computational complexity. COOLCAT is information theory based approach. It tries to optimize the global utility function based on entropy criterion. The original approach is highly dependent on the initial conditions. LIMBO is a scalable hierarchical categorical clustering algorithm built on the Information Bottleneck framework. As a hierarchical algorithm, LIMBO is not as fast as partitional methods.

In the experiment phase, we compare the results from our proposed approach with both other extensions of k-means algorithm and other approaches for categorical clustering.

3 Definitions and Notation

We assume that the set of objects to be clustered is stored in a dataset $\mathbf{D}$ defined by a set of attributes $A_1, \ldots, A_m$ with domains $\mathcal{D}_1, \ldots, \mathcal{D}_m$, respectively. Each object in $\mathbf{D}$ is represented by a tuple $t \in \mathcal{D}_1 \times \ldots \times \mathcal{D}_m$. Regarding the clustering problem discussed in this paper, we consider two types of data: *numeric* and *categorical*. The domain of numerical attributes consists of continuous real values. Thus, the distance measure based on geometric concept such as the Euclid distance or Manhattan distance can be used. A domain $\mathcal{D}_i$ is defined as categorical if it is finite and unordered, that only a comparison operation is allowed in $\mathcal{D}_i$. It means, for any $x, y \in \mathcal{D}_i$, we have either $x = y$ or $x \neq y$.

For simplicity, we represent each data object X as a tuple $(x_1, ..., x_m) \in \mathcal{D}_1 \times ... \times \mathcal{D}_m$.

4 k-means Algorithm

The k-means algorithm [2] is one of the most popular algorithm in *partitional* or *non-hierarchical* clustering methods. Given a set $\mathbf{D}$ of n numerical data objects, a natural number $k \leq n$, and a distance measure d, the k-means algorithm searches for a partition of $\mathbf{D}$ into k non-empty disjoint clusters that minimizes the overall sum of the squared distances between data objects and their cluster centers. The problem can be formulated in terms of an optimization problem as follow:

Minimize

$$P(W, \mathcal{Q}) = \sum_{l=1}^{k} \sum_{i=1}^{n} w_{i,l} d(X_i, Q_l) \qquad (1)$$

subject to

$$\sum_{l=1}^{k} w_{i,l} = 1, \quad 1 \leq i \leq n,$$
$$w_{i,l} \in \{0, 1\}, \quad 1 \leq i \leq n, 1 \leq l \leq k, \qquad (2)$$

where $W = [w_{i,l}]_{n \times k}$ is a partition matrix ($w_{i,l}$ take value 1 if object X_i is in cluster C_l, and 0

otherwise), $\mathcal{Q} = \{\mathcal{Q}_1, \ldots, \mathcal{Q}_k\}$is a set of cluster centers, and $d(.,.)$ is the squared Euclidean distance between two objects.

The problem P can be solved by iteratively solving two problems:
- Fix $\mathcal{Q} = \hat{Q}$ then solve the reduced problem $P(W, \hat{Q})$ to find $\hat{W}$.
- Fix $W = \hat{W}$ then solve the reduced problem $P(\hat{W}, Q)$.

Basically, the k-means algorithm iterates through a three-step process until $P(W, \mathcal{Q})$ converges to some local minimum:

1. Select an initial $\mathcal{Q}^{(0)} = Q_1^{(0)}, \ldots, Q_k^{(0)}$, and set $t = 0$.
2. Keep $\mathcal{Q}^{(t)}$ fixed and solve $P(W, \mathcal{Q}^{(t)})$ to obtain $W^{(t)}$, that is, having the cluster centers, we then assign each object to the cluster of its nearest cluster center.
3. Keep $W^{(t)}$ fixed and generate $\mathcal{Q}^{(t+1)}$ such that $P(W^{(t)}, \mathcal{Q}^{(t+1)})$ is minimized, that is, construct new cluster centers according to the current partition.

In the case of convergence or if a given stopping criterion is fulfilled, output the result and stop. Otherwise, set $t = t + 1$ and go to step 2.

In numerical clustering problem, the Euclidean norm is often chosen as a natural distance measure in the k-means algorithm. With this distance measure, we calculate the partition matrix in step 2 as below, and the cluster center is computed by the mean of cluster's objects.

$$\begin{aligned} &\text{if } d(X_i, C_l) \leq d(X_i, C_p) \text{ then} \\ &w_{i,l} = 1, \text{and } w_{i,p} = 0, \quad \text{for } 1 \leq p \leq k, p \neq l \end{aligned} \tag{3}$$

It was also shown in [4] that the k-means method can be extended to categorical data by using a simple matching distance measure for categorical objects with the most frequent values to define the "cluster centers" called modes. Let X, Y are two categorical objects in $\mathbf{D}$, with $X = (x_1, \ldots, x_m$ and $Y = (y_1, \ldots, y_m)$. The dissimilarity between X and Y can be computed by the total matching of the corresponding attribute values of the two objects. Formally,

$$d(X, Y) = \sum_{i=1}^{m} \delta(x_i, y_i) \tag{4}$$

where

$$\delta(x_i, y_i) = \begin{cases} 0 & \text{if } x_i = y_i, \\ 1 & \text{if } x_i \neq y_i. \end{cases}$$

Given a cluster $X_1, \ldots, X_p$ of categorical objects, with $X_i = (x_{i,1}, \ldots, x_{i,m})$, $1 \leq i \leq p$, its mode $Q = (q_1, \ldots, q_m)$ is defined by assigning q_j, $1 \leq j \leq m$, the value most frequently appeared in $\{x_{1,j}, \ldots, x_{p,j}\}$. With these modifications, Huang [5] proposed a k-modes algorithm that mimics the k-means method to cluster categorical data. However, it should be also emphasized that, by definition, the mode of a cluster is not in general unique. This makes the algorithm unstable depending on mode selection during the clustering process.

In 2004, San et al. [13] proposed a new alternative algorithm, k-representative, trying to overcome the drawback of the k-modes. The modes are replaced by "representatives". Given a cluster $C = X_1, \ldots, X_p$ of categorical objects, with

$$X_i = (x_{i,1}, \ldots, x_{i,m}), \quad 1 \leq i \leq p,$$

denote by D_j the domain set of categorical attribute A_j. Then the representative of C is defined by $Q = (q_1, \ldots, q_m)$, with

$$q_j = \{(c_j, f_{c_j}) | c_j \in D_j\} \tag{5}$$

where f_{c_j} is the relative frequency of category c_j in C, i.e., $f_{c_j} = n_{c_j}/p$, where n_{c_j} is the number of objects in C having value c_j at attribute A_j.

The dissimilarity between a categorical object $X = (x_1, \ldots, x_m)$ and the cluster representative Q is calculated based on simple matching (cf.(4)) as follows

$$d(X, Q) = \sum_{j=1}^{m} \sum_{c_j \in D_j} f_{c_j}.\delta(x_j, c_j) \tag{6}$$

The formulation shows that the dissimilarity $d(X, Q)$ is mainly dependent on the relative frequencies of categorical values within the cluster and the simple matching between categorical value.

5 Proposed Approach

5.1 Proposed dissimilarity measure

In this paper, instead of using simple matching measure, we propose a new dissimilarity measure which is based on an information-theoretic definition of similarity [15]. For example, x and y are values from categorical values set of attribute A_j. We can describe x and

y with the proposition like: value_$A_j(x)$ and value_$A_j(y)$. The proposition that states the commonality between x and y is "value_$A_j(x)$ and value_$A_j(y)$". In information theory, we can calculate the information contained in a statement by the negative logarithm of the probability of the statement. From that, the similarity between x and y is measured by the ratio between the amount of information needed to state the commonality of x and y and the information needed to fully describe what x and y are. Specifically, we define the measure as,

Let D be the set of n categorical objects defined by a set of attributes $A_1, ..., A_m$ with domains $D_1, \ldots, D_m$. For any $x, y \in D_j$,

$$\delta(x, y) = 1 - \frac{2 \times \log f_{x,y}}{\log f_x + \log f_y}, \qquad (7)$$

where f_x, f_y are the frequency of the value x and y, respectively at the attribute A_j in D. $f_{x,y}$ is the frequency that an object in D received value either x or y at attribute A_j. We can see that, if x and y are identical then $\delta(x, y) = 0$, otherwise $0 < \delta(x, y) < 1$.

5.2 k-Representatives Algorithm with proposed dissimilarity measure

With the modification above, we formulate the problem of clustering categorical data in much the same way as in the k-representative algorithm proposed in [13].

Assume that we have a data set $\mathbf{D} = X_1, \ldots, X_n$ of categorical objects to be clustered, where each object $X_i = (x_{i,1}, \ldots, x_{i,m})$, $1 \leq i \leq n$ is described by m categorical attributes. Mathematically, the problem can be stated as: Minimize

$$P(W, \mathcal{Q}) = \sum_{l=1}^{k} \sum_{i=1}^{n} w_{i,l} d(X_i, Q_l) \qquad (8)$$

subject to

$$\sum_{l=1}^{k} w_{i,l} = 1, \quad 1 \leq i \leq n,$$

$$w_{i,l} \in \{0, 1\}, \quad 1 \leq i \leq n, 1 \leq l \leq k, \qquad (9)$$

where $W = [w_{i,l}]_{n \times k}$ is a partition matrix, $\mathcal{Q} = \{\mathcal{Q}_1, \ldots, \mathcal{Q}_k\}$is a set of cluster representatives, and $d(.,.)$ is the dissimilarity measure defined by (6) with the modified $\delta(.,.)$ in (7). The algorithm is processed as:

1. Randomly initialize a k-partition of $\mathbf{D}$.
2. Calculate the representatives for each cluster.
3. Reassign objects to clusters regarding the dissimilarity between objects and cluster representatives. When finish, recalculate the representatives.
4. Repeat step 3 until the partition has no change or the number of iterations is beyond user-defined threshold.

6 Experiments and Results

In this section we present a comprehensive evaluation of our approach. In terms of extension of k-means algorithm for categorical objects, our approach is compared with k-representative algorithm. And we also want to compare the performance of our approach with other methods for clustering categorical data such as: ROCK [10] and LIMBO [11].

6.1 Datasets

For the evaluation, we used real world data sets downloaded from the UCI Machine Learning Repository [18]. The main characteristics of the datasets are summarized in Table 1. These datasets are chosen to test our algorithm because of their public availability and since all attributes can be treated as categorical ones.

Datasets	Number of Instances	Number of Attributes	Number of Classes
Car	1728	6	4
Soybean (Small)	47	35	4
Soybean (Large)	683	35	19
Votes	435	16	2
Nursery	12960	8	5
Mushroom	8124	22	2

Table 1: Datasets Summarization

6.2 Clustering Evaluation Metrics

Evaluating the clustering quality is often a hard and subjective task [16]. Generally, objective functions in clustering are purposely designed so as to archive high intra-cluster similarity and low inter-cluster similarity. This can be viewed as an internal criterion for the quality of a clustering. However, as observed in the literature, good

scores on an internal criterion do not necessarily translate into good effectiveness in an application. Here, by the same way as in [17], we use three external criteria to evaluate the results: Purity, Normalized Mutual Information (NMI) and Adjusted Rand Index (ARI). These methods make use of the original class information of each object and the cluster to which the same objects have been assigned to evaluate how well the clustering result matches the original classes.

We denote by $C = \{C_1, \ldots, C_J\}$ the partition of the dataset built by the clustering algorithm, and by $P = \{P_1, \ldots, P_I\}$ the partition inferred by the original classification. J and I are respectively the number of clusters $|C|$ and the number of classes $|P|$. We denote by n the total number of objects.

Purity is a simple and transparent evaluation measure. To compute purity , each cluster is assigned to the class which is most frequent in the cluster, and then the accuracy of this assignment is measured by counting the number of correctly assigned objects and dividing by the number of objects in the dataset. High purity is easy to achieve when the number of clusters is large. Thus, we cannot use purity to trade off the quality of the clustering against the number of clusters.

$$Purity(C, P) = \frac{1}{n} \sum_j \max_i |C_j \cap P_i| \quad (10)$$

The second metric (NMI) provides an information that is independent from the number of clusters [19]. This measure takes its maximum value when the clustering partition matches completely the original partition. NMI is computed as the average mutual information between any pair of clusters and classes

$$NMI(C, P) = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} |C_j \cap P_i| \log \frac{n|C_j \cap P_i|}{|C_j||P_i|}}{\sqrt{\sum_{j=1}^{J} |C_j| \log \frac{|C_j|}{n} \sum_{i=1}^{I} |P_i| \log \frac{|P_i|}{n}}} \quad (11)$$

The third metric is the adjusted Rand index [20]. Let a be the number of object pairs belonging to the same cluster in C and to the same class in P. This metric captures the deviation of a from its expected value corresponding to the hypothetical value of a obtained when C and P are two random, independent partitions. The expected value of a denoted by $E[a]$ is computed as follows:

$$E[a] = \frac{\pi(C)\pi(P)}{n(n-1)/2} \quad (12)$$

where $\pi(C)$ and $\pi(P)$ denote respectively the number of object pairs from the same clusters in C and from the same class in P. The maximum value for a is defined as:

$$\max(a) = \frac{1}{2}(\pi(C) + \pi(P)) \quad (13)$$

The agreement between C and P can be estimated by the adjusted rand index as follows:

$$ARI(C, P) = \frac{a - E[a]}{\max(a) - E[a]} \quad (14)$$

when $ARI(C, P) = 1$, we have identical partitions.

6.3 Results

We run the experiments on a Mac with a 3.66GHz Intel QuadCore processor, 8GB of RAM running Mac OSX 10.8 . The performance of three evaluation metrics are calculated by the average after running the algorithm 1000 times on each dataset. For the performance result of ROCK and LIMBO, we obtained from the experiments in [17]. In the tables 2, 3 and 4 we report the results of the comparative evaluation against ROCK, LIMBO and k-representative. From the results, we can see that our approach and k-representative are very much close to each other in terms of performances with almost datasets, except the nursery dataset where our approach obtain slightly better results in all three criteria. It can be seen that, ROCK and LIMBO are not stable enough. In some case their performance are good enough, but in other case the performance are far from the results of our approach. For example, ROCK is worst with mushroom dataset but quite good with car dataset or vote dataset. And LIMBO are worst with car dataset but has best results with mushroom dataset. On the other hand, our approach performances are quite stable. Regarding the purity criteria, we got the best results on 3 out of 4 datasets. And for two remain metrics, we also got best results in 2 out of 4 datasets and the other one is close to the best. Especially, with car dataset, our modified k-representative algorithm outperform ROCK and LIMBO in all evaluation metrics.

7 Conclusion

In this paper, we proposed a new approach to define the dissimilarity for categorical data and

Datasets	Proposed algorithm	k-representative	ROCK	LIMBO
Car	0.71	0.7	0.7	0.44
Soybean (Small)	0.93	0.93	N/A	N/A
Soybean (Large)	0.56	0.57	0.44	0.56
Votes	0.87	0.88	0.84	0.87
Nursery	0.46	0.44	N/A	N/A
Mushroom	0.78	0.77	0.5	0.89

Table 2: Purity results

Datasets	Propsed algorithm	k-representative	ROCK	LIMBO
Car	0.12	0.11	0.036	0.014
Soybean (Small)	0.93	0.93	N/A	N/A
Soybean (Large)	0.7	0.7	0.63	0.75
Votes	0.5	0.49	0.34	0.44
Nursery	0.11	0.094	N/A	N/A
Mushroom	0.35	0.35	0.057	0.57

Table 3: NMI results

applied the method to an alternative algorithm that mimics the k-means method in clustering categorical data. The performance of new approach is compared with other categorical clustering algorithms. The experimental results have shown that the proposed approach gets good results and is more stable than ROCK and LIMBO algorithm. In addition, it is also comparable or slightly better to the original k-representative algorithm. However, the improvement is less than what we expected. As for future work, we would like to apply another way to assign the dissimilarity between two matches of categorical attribute values (in the current measure, matches are uniformly given the maximum value as 1). For example, we could give lower weight to matches on frequent values, and higher weight to matches on infrequent values or rare values. It is also of interest to extend this technique to the problem of clustering for mixed datasets as well as to the problem of fuzzy clustering for categorical data.

References

[1] J. Han, M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, San Francisco, 2001.

[2] J.B. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Symposium on Mathematical Statistics and Probability*, Berkelely, CA, 1967, vol. 1, no. AD 669871, 281–297.

[3] S.Z. Selim, M.A. Ismail. k-Means-type algorithms: A generalized convergence theorem and characterization of local optimality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6: 81–87, 1984.

[4] Z. Huang. Clustering large data sets with mixed numeric and categorical values. In H. Lu, H. Motoda and H. Luu, editors, *KDD: Techniques and Applications*, pages 21–34. World Scientific, Singapore, 1997.

[5] Z. Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, 2: 283–304, 1998.

[6] Z. Huang, Michael K. Ng. A fuzzy k-modes algorithm for clustering categorical data. *IEEE Transactions on Fuzzy Systems*, 7: 446–452, 1999.

[7] D. Gibson, J. Kleinberg, P. Raghavan. Clustering categorical data: An approach based on dynamic systems. In *Proceedings of the 24-th International Conference on Very Large Databases*, New York, 1998, pp. 311–323.

[8] S. Guha, R. Rastogi and K. Shim. CURE: An efficient clustering algorithm for large databases. In *Proceedings of ACM SIGMOD International Conference on Management of Data*, New York, 1998, pp. 73–84.

[9] V. Ganti, J. Gehrke, Raghu Ramakrishnan. CATUS - Clustering categorical data using

Datasets	Proposed algorithm	k-representative	ROCK	LIMBO
Car	0.06	0.052	0.004	0.037
Soybean (Small)	0.88	0.88	N/A	N/A
Soybean (Large)	0.38	0.37	0.17	0.44
Votes	0.56	0.56	0.46	0.55
Nursery	0.075	0.058	N/A	N/A
Mushroom	0.39	0.39	-0.001	0.61

Table 4: Adjusted Rand Index results

summaries. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, (San Diego, USA), pp. 73–83, 1999.

[10] S. Guha, R. Rastogi, K. Shim. ROCK: A robust clustering algorithm for categorical attributes, *Information Systems* **25**(5) (2000) 345–366.

[11] P. Andritsos, P. Tsaparas, R.J. Miller, K.C. Sevcik. *LIMBO: A Scalable Algorithm to Cluster Categorical Data* 2003

[12] D. Barbara, J. Couto, Y. Li. COOLCAT: an entropy-based algorithm for categorical clustering In *Proceedings of the eleventh international conference on Information and knowledge management*, 582–589, 2002.

[13] O.M. San, V.-N. Huynh, Y. Nakamori. An alternative extension of the k-means algorithm for clustering categorical data *Int. J. Appl. Math. Comput. Sci*, pp 241–247, 2004.

[14] M. K. Ng, M. J. Li, J. Z. Huang, Z. He. On the Impact of Dissimilarity Measure in k-Modes Clustering Algorithm *IEEE Trans. Pattern Anal. Mach. Intell.* **29** (2007) 503–507.

[15] D. Lin. An Information-Theoretic Definition of Similarity In*Proceedings of the 15th International Conference on Machine Learning*, pp 296–304, 1998.

[16] D. Ienco, R.G. Pensa, R. Meo. Context-based Distance Learning for CategoricalData Clustering *Advances in Intelligent Data Analysis VIII: 8th International Symposium*, Springer, pp. 83–94, 2009.

[17] D. Ienco, R.G. Pensa, R. Meo. From Context to Distance: Learning Dissimilarity for Categorical Data Clustering *ACM Trans. Knowl. Discov. Data* New York, 2012.

[18] Blake, C.L. and C.J. Merz. UCI Repository of machine learning databases [http://www.ics.uci.edu/˜mlearn/MLRepository.html]. Irvine, CA: University of California, Department of Information and Computer Science, 1998.

[19] Strehl, A., Ghosh, J., and Cardie, C. Cluster ensembles - a knowledge reuse framework for combining multiple partitions *Journal of Machine Learning Research 3*, 583–617, 2002.

[20] Hubert, L. and Arabie, P Comparing partitions *Journal of Classification 2*, 1, 193–218, 1995.

A Computational Method for Obtaining Stackelberg Solutions to Two-level Nonlinear Programming Problems through Strategic Oscillation Using Particle Swarm Optimization

Kosuke Kato[†] **Shimpei Matsumoto**[†] **Shuhei Mimura**[†]
Nobuyuki Yamawaki[‡] **Hideki Katagiri**[*]
[†]Faculty of Applied Information Science, Hiroshima Institute of Technology, JAPAN
[‡]Faculty of Biology-Oriented Science and Technology, Kinki University, JAPAN
[*]Graduate School of Engineering, Hiroshima University, JAPAN
{k.katoh.me, s.matsumoto.gk}@it-hiroshima.ac.jp

Abstract

In economical or public decision making situations, there often exist two decision makers. In particular, the situation where one of them has the priority in decision over the other is mathematically defined as two-level programming problems. When it is assumed that each of them can obtain the information of the objective function and constraints for both itself and the other and does not have motivation to cooperate mutually, the Stackelberg solution is adopted as a reasonable solution. Since problems to obtain Stackelberg solutions are generally nonconvex programming problems with complex structures, it is not easy to find exact Stackelberg solutions. In the present paper, we try to obtain (approximate) Stackelberg solutions for two-level programming problems through strategic oscillation using particle swarm optimization (PSO).

Keywords: Two-level programming, Stackelberg solution, strategic oscillation, particle swarm optimization

1 Introduction

In economical or public decision making situations, there often exist two decision makers. In particular, the situation where one of them has the priority in decision over the other is mathematically defined as two-level programming problems. In two-level linear programming problems, a decision maker with the priority in decision is called the upper level decision maker and the other is called the lower level decision maker. Each of them has its own decision variables and objective function. Two major approaches for two-level programming problems are adopted depending on the situations which the DMs are placed in. One approach is considered under the assumption that decision makers can cooperate with each other. For instance, the upper level decision maker is the administrative office of a company and the lower level decision maker is one of its autonomous divisions. Lai [1] and Shih et al. [2] proposed fuzzy programming approaches. Abo-Sinna et al. [3] discussed an interactive approach for multi-level multi-objective programming problems.

On the other hand, when it is assumed that each of decision makers can obtain the information of the objective function and constraints for the other as well as itself and does not have motivation to cooperate mutually, the Stackelberg solution [4] is adopted as a reasonable solution. Since problems to obtain Stackelberg solutions are generally nonconvex programming problems with complex structures, it is not easy to find exact Stackelberg solutions. For two-level non-linear programming problems, Gümüs et al. [5] proposed an exact solution method. Since it needs the differentiability of objective functions and constraint ones, it cannot be applied to two-level nonlinear programming problems with nondifferentiable objective functions or constraint functions. Thus, as efficient computational methods, meraheuristics are considered to be efficient, e.g., genetic algorithms [6], particle swarm optimization [7], evolutionary multi-agent systems [8]. The effectiveness of these methods for finding (approximate) Stackelberg solutions to two-level nonlinear programming problems are shown in those papers. However, there remains a problem that they are time-consuming since these methods deal with so many individuals by adopting hierarchical populations.

In the present paper, we try to obtain (approximate) Stackelberg solutions for two-level nonlinear programming problems through strategic oscillation using particle swarm optimization (PSO) [9] without hierarchical populations.

2 Two-level nonlinear programming problems

In economical or public decision making problems, generally, there exist plural decision makers who has independent concerns. In this paper, as a mathematical model of a decision making problem where each of decision makers has its own objective function conflicting with that of the other and makes a decision independently and sequentially, we consider two-level nonlinear programming problems formulated as:

$$\left.\begin{array}{l} \underset{\text{DM1 (upper level)}}{\text{minimize}}\ f_1(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \underset{\text{DM2 (lower level)}}{\text{minimize}}\ f_2(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \quad \text{subject to } g_i(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0, \ \ i = 1, 2, \ldots, m \end{array}\right\} \tag{1}$$

where $\boldsymbol{x}_1$ is an n_1 dimensional decision variable column vector for the upper level decision maker (DM1), $\boldsymbol{x}_2$ is an n_2 dimensional decision variable column vector for the lower level decision maker (DM2), $f_1(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the objective function for DM1, $f_2(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the objective function for DM2 and $g_i(\boldsymbol{x}_1, \boldsymbol{x}_2)$, $i = 1, 2, \ldots, m$ are constraint functions. In general, $f_l(\cdot)$, $l = 1, 2$ and $g_i(\cdot)$, $i = 1, 2, \ldots, m$ are nonlinear.

In this paper, we consider the case that decision makers in (1) do not have motivation to cooperate mutually, i.e., they are uncooperative. Under the uncooperative assumption, the Stackelberg solution is generally adopted as the concept of a reasonable solution to (1), Stackelberg solution is an optimal solution for the upper level decision maker DM1 under the assumption that both decision makers know the opponent's objective function and constraint functions, and the lower level decision maker DM2 makes a decision to optimize its own objective corresponding to a given decision by DM1. Namely, the Stackelberg solution is an optimal solution to the following problem.

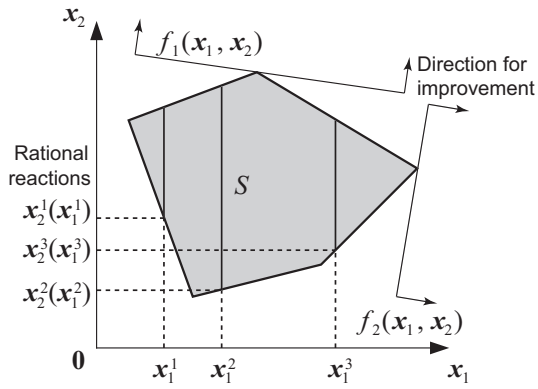


Figure 1. An example of rational reactions

$$\left.\begin{array}{l} \underset{\boldsymbol{x}_1}{\text{minimize}}\ f_1(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \text{where } \boldsymbol{x}_2 \text{ solves} \\ \quad \underset{\boldsymbol{x}_2}{\text{minimize}}\ f_2(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \quad \text{subject to } g_i(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0, \ \ i = 1, 2, \ldots, m \end{array}\right\} \tag{2}$$

Let the feasible region of (2) be denoted by S. To be more specific, if a feasible $\boldsymbol{x}_1$ (a decision of DM1) in S is given, it is assumed that the lower level decision maker DM2 selects an element $\boldsymbol{x}_2(\boldsymbol{x}_1)$ in a set $R(\boldsymbol{x}_1)$ which consists of optimal solutions to the following problem to optimize its own objective function $f_2(\cdot)$

$$\left.\begin{array}{l} \underset{\boldsymbol{x}_2}{\text{minimize}}\ f_2(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \text{subject to } g_i(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0, \ \ i = 1, 2, \ldots, m \end{array}\right\} \tag{3}$$

where an optimal solution to (3) is called a rational reaction to $\boldsymbol{x}_1$ (Fig. 1).

Therefore, the upper level decision maker DM1 reasonably selects a decision $\boldsymbol{x}_1$ which optimizes its own objective function $f_1(\cdot)$ from the inducible region $IR = \{(\boldsymbol{x}_1, \boldsymbol{x}_2) \mid (\boldsymbol{x}_1, \boldsymbol{x}_2) \in S, \boldsymbol{x}_2 \in R(\boldsymbol{x}_1)\}$, where the inducible region is the set of rational reactions for all feasible $\boldsymbol{x}_1$. This decision by DM1 is called Stackelberg solution which is obtained by solving following problem

$$\left.\begin{array}{l} \underset{\boldsymbol{x}_1}{\text{minimize}}\ f_1(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ \text{subject to } g_i(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0, \ \ i = 1, 2, \ldots, m \\ \qquad\qquad \boldsymbol{x}_2 \in R(\boldsymbol{x}_1) \end{array}\right\} \tag{4}$$

An example of the inducible region and Stackelberg solution is illustrated in Fig. 2.

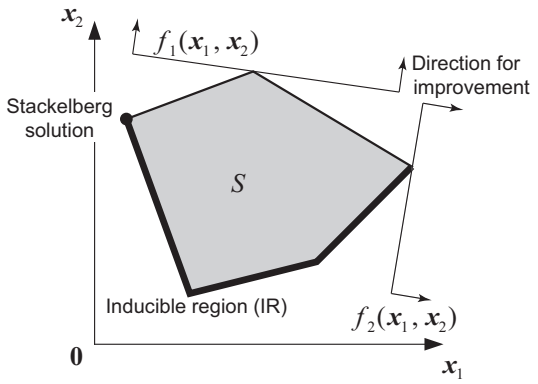


Figure 2. An example of the inducible region and Stackelberg solution

3 Derivation of Stackelberg solution by strategic oscillation and PSO

Hanafi and Freville [10] proposed an efficient tabu search through strategic oscillation for multidimensional 0-1 knapsack problems. Focusing on the fact that an optimal solution to any multidimensional 0-1 knapsack problem exists on the boundary of the feasible region, they regarded the area around the boundary of the feasible region as the promising zone where the optimal solution must exist.

In the tabu search, for the purpose of the global search, they adopted the strategic oscillation which consists of two phases: one phase is the intensification where the promising zone is intensively searched, the other is the diversification where the structure of a solution is drastically changed by intentionally moving it to the internal or external area of the feasible region (Fig. 3).

In this research, noting that Stackelberg solutions to two-level programming problems necessarily exist in the inducible region IR, we propose a new computational method to obtain (approximate) Stackelberg solution through strategic oscillation using particle swarm optimization (PSO) [9] where the inducible region IR is regarded as the promising zone.

Proposed algorithm

Step 1: Move from the internal area of the feasible region (FR) to the promising zone (PZ)
For the current search point (solution) $(\boldsymbol{x}_1, \boldsymbol{x}_2)$, let $\boldsymbol{x}_1$ be fixed, $\boldsymbol{x}_2$ be variable. Next, find the (approximate) rational reaction $\boldsymbol{x}_2(\boldsymbol{x}_1)$ in the sense of minimizing $f_2(\cdot)$ under the constraints using particle swarm optimization (PSO). Then, update the search point to $(\boldsymbol{x}'_1, \boldsymbol{x}'_2) := (\boldsymbol{x}_1, \boldsymbol{x}_2(\boldsymbol{x}_1))$. Go to step 2.

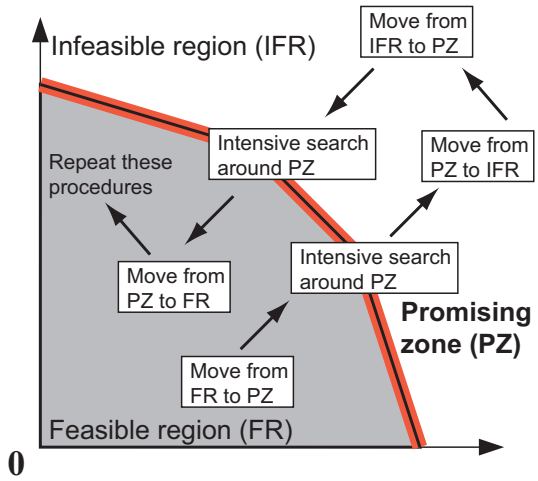


Figure 3. The strategic oscillation for multidimensional 0-1 knapsack problem

Step 2: Intensive search the promising zone (PZ) around the current search point
For the current search point $(\boldsymbol{x}'_1, \boldsymbol{x}'_2)$, let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be variable. Next, find the best search point in the sense of minimizing $f_1(\cdot)$ under the constraints where $\boldsymbol{x}_2$ is restricted around the current $\boldsymbol{x}'_2$ using PSO. Then, let the obtained best search point be denoted by $(\boldsymbol{x}''_1, \boldsymbol{x}''_2)$. If the termination condition is fulfilled, stop. Otherwise, go to step 3.

Step 3: Search of the extended promising zone including the infeasible region (IFR)
For the current search point $(\boldsymbol{x}''_1, \boldsymbol{x}''_2)$, let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be variable. Next, find the best search point in the sense of minimizing $f_1(\cdot)$ without considering the constraints for a given number of times using PSO. Then, let the obtained best search point be denoted by $(\boldsymbol{x}'''_1, \boldsymbol{x}'''_2)$. Go to step 4.

Step 4: Move from the extended promising zone to the promising zone (PZ)
For the current search point $(\boldsymbol{x}'''_1, \boldsymbol{x}'''_2)$, let $\boldsymbol{x}_1$ be fixed, $\boldsymbol{x}_2$ be variable. Next, find a feasible solution using PSO by changing $\boldsymbol{x}_2$.

If we cannot find any feasible solution by changing $\boldsymbol{x}_2$, let $\boldsymbol{x}_1$ be variable and search a feasible solution by changing $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. After finding a feasible solution, let $\boldsymbol{x}_1$ be fixed, $\boldsymbol{x}_2$ be variable. Then, search the (approximate) rational reaction $\boldsymbol{x}_2(\boldsymbol{x}_1)$ in the sense of minimizing $f_2(\cdot)$ under the constraints using particle swarm optimization (PSO). Through the above procedure, obtain the (approximate) rational reaction and let the obtained search point be denoted by $(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*)$. Go to step 5.

Step 5: Intensive search the promising zone (PZ) around the current search point
For the current search point $(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*)$, carry out the same procedure in step 2. Then, let the obtained best search point be denoted by $(\boldsymbol{x}_1^{**}, \boldsymbol{x}_2^{**})$. Go to step 6.

Step 6: Move from the promising zone (PZ) to the internal area of the feasible region (FR)
For the current search point $(\boldsymbol{x}_1^{**}, \boldsymbol{x}_2^{**})$, let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be variable. Next, find the best search point in the sense of maximizing the surplus of constraints for a given number of times using PSO. Then, let the obtained best search point be denoted by $(\boldsymbol{x}_1^{***}, \boldsymbol{x}_2^{***})$. Go to step 1.

Here, we brief the overview of PSO used in the above algorithm. Particle swarm optimization (PSO) [9] is based on the social behavior that a population of individuals adapts to its environment. This adaptation to the environment is a stochastic process that depends on both the memory of each individual, called particle, and the knowledge gained by the population, called swarm.

In the implementation of PSO, a particle has four attributes: the current position vector in the search space, the current velocity vector, the best position in its track and the best position of the swarm. The algorithm is outlined as follows.

Step 1: Generate the initial swarm involving N particles randomly.

Step 2: Calculate the new velocity vector of each particle, based on its attributes.

Step 3: Calculate the new position of each particle from the current position vector and its new velocity vector.

Step 4: If the termination condition is fulfilled, stop. Otherwise, go to Step 2.

More precisely, the new velocity vector of the i-th particle at time t, $\boldsymbol{v}_i^{t+1}$, is calculated by the following scheme [11].

$$\boldsymbol{v}_i^{t+1} := \omega^t \boldsymbol{v}_i^t + c_1 R_1^t (\boldsymbol{p}_i^t - \boldsymbol{x}_i^t) + c_2 R_2^t (\boldsymbol{p}_g^t - \boldsymbol{x}_i^t) \tag{5}$$

In (5), R_1^t and R_2^t are randomly determined on the interval $[0, 1)$, $\boldsymbol{p}_i^t$ is the best position of the i-th particle in its track and $\boldsymbol{p}_g^t$ is the best position of the swarm. Parameters c_1 and c_2 are called trust parameters. In this paper, they are set to 2.0. Another parameter ω^t is called the inertia parameter of the particle (the value of ω^t is explained below).

Then, calculate the new position of the i-th particle at time t, $\boldsymbol{x}_i^{t+1}$, according to (6).

$$\boldsymbol{x}_i^{t+1} := \boldsymbol{x}_i^t + \boldsymbol{v}_i^{t+1} \tag{6}$$

where $\boldsymbol{x}_i^t$ is the current position of the i-th particle at time t.

Namely, after calculating the next search direction vector $\boldsymbol{v}_i^{t+1}$ for the i-th particle by (5) on the basis of the current search direction vector $\boldsymbol{v}_i^t$, the direction vector from the current search position $\boldsymbol{x}_i^t$ to the best position in its track $\boldsymbol{p}_i^t$ and the direction vector from the current search position $\boldsymbol{x}_i^t$ to the best position of the swarm $\boldsymbol{p}_g^t$, calculate the next search position $\boldsymbol{x}_i^{t+1}$ by (6) based on the current position $\boldsymbol{x}_i^t$ and the next search direction vector $\boldsymbol{v}_i^{t+1}$.

The parameter ω^t controls the amount of the move to search globally in early stage and to search locally by decreasing ω^t gradually. It is defined by (7)

$$\omega^t := \begin{cases} \omega^0 - \dfrac{t(\omega^0 - \omega^{T_{\max}})}{0.75 \times T_{\max}} & , t \leq 0.75 \times T_{\max} \\ \omega^{T_{\max}} & , t > 0.75 \times T_{\max} \end{cases} \tag{7}$$

where $T_{\max}$ is the number of maximum iteration times, ω^0 is an initial value at the time iteration, $\omega^{T_{\max}}$ is the last value. The searching procedure of PSO is illustrated in Fig. 4. In this paper, the constraint of problem is handled by adding the penalty term to the objective function.

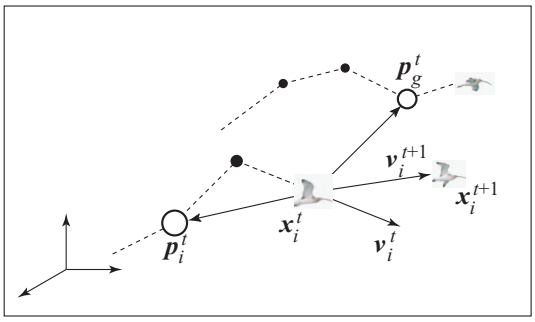


Figure 4. Movement of a particle in PSO.

Table 1. Values of upper level objective functions f_1 and average computational times (10 trials).

	Proposed
Best of f_1	1.25000
Average of f_1	1.25000
Worst of f_1	1.25000
x_1 for Best	5.5000
x_2 for Best	1.5000
Average time [sec]	0.65

4 Numerical examples

In order to investigate the effectiveness of the proposed method, we apply the proposed method to numerical examples. Our computational environment is summarized as CPU : Core i5 2450Hz, Memory : 8GB, OS: Windows 7, Programming Language: C.

First, we apply the proposed method to the following simple two-level nonlinear programming problem.

$$\left.\begin{array}{l}
\underset{x_1}{\text{minimize}}\ f_1(x_1,x_2) = (x_1-5)^2 + (2x_2-2)^2 \\
\text{where } x_2 \text{ solves} \\
\quad \underset{x_2}{\text{minimize}}\ f_2(x_1,x_2) = x_1^3 - 1.5x_1x_2 + (x_2-1)^2 \\
\quad \text{subject to } -3x_1 + x_2 + 3 \le 0 \\
\qquad\qquad\qquad x_1 - x_2 - 4 \le 0 \\
\qquad\qquad\qquad x_1 + x_2 - 7 \le 0 \\
\qquad\qquad x_1 \ge 0,\ \ x_2 \ge 0
\end{array}\right\} \tag{8}$$

The Stackelberg solution to (8) is $(x_1, x_2) = (11/2, 3/2)$ and its upper level objective function value $f_1(x_1, x_2)$ is $5/4$ (Fig. 5).

The results are shown in table 1. Table 1 indicates that the proposed method succeeds in finding the exact Stackelberg solution in 1 second.

Next, we apply the proposed method to the following two-level nonlinear programming problem which is similar to (8) and compare the result with those for two existing methods (PSO[7], EMAS[8]) reported in [7].

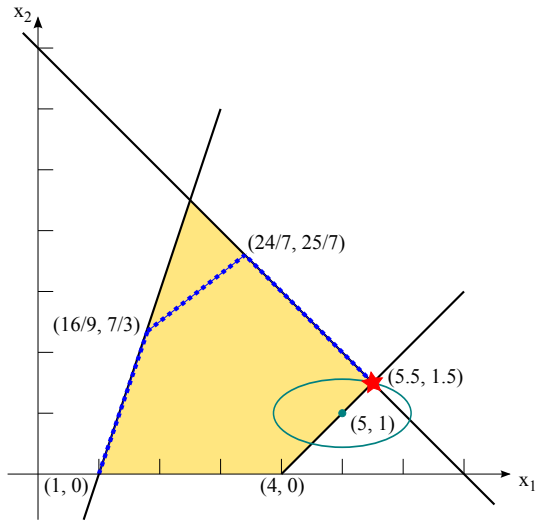


Figure 5. The inducible region (dotted line) and Stackelberg solution (star) for problem (8)

$$\left.\begin{array}{l}
\underset{x_1}{\text{minimize}}\ f_1(x_1,x_2) = (x_1-5)^2 + (2x_2+1)^2 \\
\text{where } x_2 \text{ solves} \\
\quad \underset{x_2}{\text{minimize}}\ f_2(x_1,x_2) = x_1^3 - 1.5x_1x_2 + (x_2-1)^2 \\
\quad \text{subject to } -3x_1 + \quad x_2 + 3 \le 0 \\
\qquad\qquad\qquad x_1 - 0.5x_2 - 4 \le 0 \\
\qquad\qquad\qquad x_1 + \quad x_2 - 7 \le 0 \\
\qquad\qquad x_1 \ge 0,\ \ x_2 \ge 0
\end{array}\right\} \tag{9}$$

The Stackelberg solution to 9 is $(x_1, x_2) = (1, 0)$ and its upper level objective function value $f_1(x_1, x_2)$ is 17 (Fig. 6).

The results are shown in table 2. Table 2 indicates that both existing method can find the exact Stackelberg solution but the proposed method cannot. However, the proposed method succeed in finding approximate Stackelberg solutions in much shorter computational time than existing methods (the computational environment for existing methods are unknown).

Furthermore, we apply the proposed method to the following two-level nonlinear programming problem and compare the result with those for two existing methods (PSO[7], EMAS[8]) re-

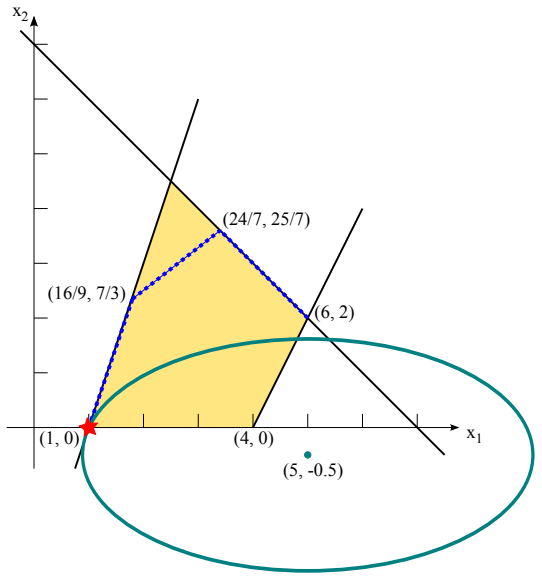


Figure 6. The inducible region (dotted line) and Stackelberg solution (star) for problem (9)

Table 2. Values of upper level objective functions f_1 and average computational times (10 trials).

	Proposed	PSO[7]	EMAS[8]
Best of f_1	17.45235	17.0000	17.0000
Average of f_1	17.45235	17.0000	17.0000
Worst of f_1	17.45235	17.0000	17.0000
x_1 for Best	1.00560	n/a	n/a
x_2 for Best	0.11178	n/a	n/a
Average time [sec]	0.71	10.11	12.33

ported in [7].

$$\left.\begin{array}{l} \underset{x_1}{\text{minimize}}\ f_1(x_1,x_2) = (x_1-5)^4 + (2x_2+1)^4 \\ \text{where } x_2 \text{ solves} \\ \underset{x_2}{\text{minimize}}\ f_2(x_1,x_2) = \exp(-x_1+x_2) + x_1^2 \\ \qquad\qquad +2x_1x_2 + x_2^2 + 2x_1 + 6x_2 \\ \text{subject to}\quad x_1 + x_2 - 4 \le 0 \\ \qquad\qquad -x_1 + x_2 - 2 \le 0 \\ \qquad\qquad x_1 \ge 0,\ \ x_2 \ge 0 \end{array}\right\} \tag{10}$$

The Stackelberg solution to 10 is $(x_1, x_2) = (4, 0)$ and its upper level objective function value $f_1(x_1, x_2)$ is 2 (Fig. 7).

The results are shown in table 3. Table 3 indicates that both existing method can find the exact Stackelberg solution but the proposed method cannot. However, the proposed method succeeds in finding approximate (almost exact) Stackelberg solutions in much shorter computational time than existing methods.

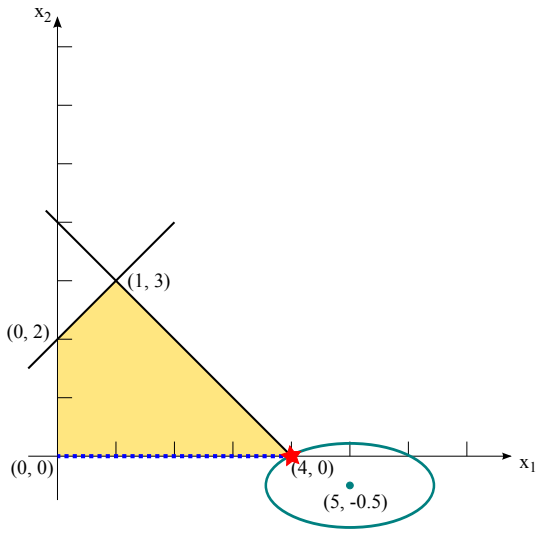


Figure 7. The inducible region (dotted line) and Stackelberg solution (star) for problem (10)

Table 3. Values of upper level objective functions f_1 and average computational times (10 trials).

	Proposed	PSO[7]	EMAS[8]
Best of f_1	2.005648	2.0000	2.0000
Average of f_1	2.005648	2.0000	2.0000
Worst of f_1	2.005648	2.0000	2.0000
x_1 for Best	3.99221	n/a	n/a
x_2 for Best	0.00011	n/a	n/a
Average time [sec]	0.57	11.45	13.29

From these results, it is claimed that the proposed method is faster than existing methods for deriving (approximate) Stackelberg solutions for two-level nonlinear programming problems. On the other hand, in the sense of solution accuracy, the proposed method is inferior to existing methods.

5 Conclusion

In this paper, we proposed a fast computational method to obtain Stackelberg solutions to noncooperative two-level nonlinear programming problems. In particular, by adopting strategic oscillation, it made the global search of the inducible region possible. In addition, by using particle swarm optimization, it made the search procedure efficient. Furthermore, we showed the efficiency of the proposed method through its applications to numerical examples. In near future,

we are going to investigate the performance of the proposed method for larger-scale problems in the real world and improve it to obtain more accurate and precise Stackelberg solutions.

References

[1] Y.J. Lai. Hierarchical optimization: a satisfactory solution. *Fuzzy Sets and Systems*, 77(3):321–335, 1996.

[2] Y.J. Lai H.S. Shih and E.S. Lee. Fuzzy approach for multi-level programming problems. *Computers & Operations Research*, 23(1):73–91, 1996.

[3] M.A. Abo-Sinna and I.A. Baky. Interactive balance space approach for solving multi-level multi-objective programming problems. *Information Sciences*, 177(16):3397–3410, 2007.

[4] Y. Ishizuka K. Shimizu and J.F. Bard. *Nondifferentiable and Two-Level Mathematical Programming*. Kluwer Academic Publishers, Boston, 1997.

[5] Z.H. Gümüs and C.A. Floudas. Global optimization of nonlinear bilevel programming problems. *Journal of Global Optimization*, 20(1):1–31, 2000.

[6] I. Nishizaki K. Niwa and M. Sakawa. Computational methods for obtaining stackelberg solutions to two-level nonlinear programming problems. *Proceedings of Second Asia-Pacific Conference on Industrial Engineering and Management Systems*, pages 489–492, 1999.

[7] M. Sakawa T. Matsui and K. Ishimaru. A computational method for obtaining stackelberg solutions to two-level nonlinear programming problems through particle swarm optimization. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, 23(3):350–362, 2011 (in Japanese).

[8] T. Matsui K. Kato, M. Sakawa and H. Ohtsuka. A computational method for obtaining stackelberg solutions to noncooperative two-level programming problems through evolutionary multi-agent systems. *Agent and Multi-Agent Systems: Technologies and Applications: Third KES International Symposium, KES-AMSTA 2009 Proceedings (Lecture Notes in Computer Science / Lecture Notes in Artificial Intelligence)*, pages 639–648, 2009.

[9] J. Kennedy and R.C. Eberhart. Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, pages 1942–1948, 1995.

[10] S. Hanafi and A. Freville. An efficient tabu search approach for the 0-1 multidimensional knapsack problem. *European Journal of Operational Research*, 106(2–3):659–675, 1998.

[11] Y.H. Shi and R.C. Eberhart. A modified particle swarm optimizer. *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 69–73, 1998.

Identifying Co-Creation Process in Energy Aggregation Services through Soft Systems Engineering

Motohisa Funabashi†

†Japan Advanced Institute of Science and Technology
funa@jaist.ac.jp

Abstract

The present paper identifies a co-creation process in energy aggregation service systems and shows future challenges for service systems engineering. First, requirements for smart energy management in local community are described and service systems engineering methodology for planning the management systems is proposed. Second, an energy aggregation service system is derived based on the methodology. Investigating the result as well as real world aggregation service systems, importance of paying attention on co-creation processes is addressed. Finally, future work for designing the co-creation process is discussed from the view of trasdisiciplinary science and technology.

Keywords: Service Co-Creation, Energy Aggregation Services, Soft Systems Methodology, Transdisciplinary Science and Technology

1 Introduction

There has been continued academic and industrial interest in the nature of services and in how to effectively market and design them since their economic importance has been recognized in 1970s. The engineering domain has also noted their importance and has been tackling how to devise methodologies for service design since 1990s. Business process reengineering began service systems engineering from the information systems engineering viewpoint.

Following this trend of business process reengineering with information and communication technology (ICT), which could be applied to all industries, Chesbrough & Spohrer [1] coined the term Service Science. They focused their attention on interactions between service providers and customers, and pointed out that interactions embody a co-creation process that includes communication through tacit knowledge. Vargo & Lusch [2] also proposed a new perspective from the marketing viewpoint for value exchanges between providers and customers as Service Dominant Logic (SDL) comparing it to Goods Dominant Logic (GDL) where people exchange to acquire the benefits of specialized competencies (knowledge and skills) or services and customers are co-producers of services. Service definitions have been moving from the transfer of value to the co-creation of value. Understanding and designing the co-creation process is essential for service science and engineering.

However so far practical very few studies on co-creation processes have been presented and discussed. The present paper deals with a co-creation process which will be required in regional energy management systems.

The organization of the paper is as follows: First, requirements and environment for smart energy management systems for local community are presented. Second, a method for deriving a service system satisfying the requirements is proposed based on soft systems methodology and an energy aggregation service system is derived by the proposed method. Third, by investigating the result and real world aggregation services, a co-creation process is identified and a possible approach for designing the co-creation process is proposed.

2 Energy Management System for Regional Community

Significant growth of global population and economic expansion of developing countries demand explosive amount of energy. The source of energy and capacity of earth for green house gas (GHG) caused by energy consumption is definitely limited. In order to achieve the sustainable society, we have to reduce energy consumptions and switch energy resources to re-

newable ones.
Of all the plans to reduce energy consumptions and GHG emissions, regional energy aggregation services are the most promising because they pursue local energy production and consumption that can reduce the loss of transmitting energy as well as appeal to people to become more energy-aware. The basic idea behind a regional energy aggregation service is outlined in Figure 1. The aggregation service provider bridges regional energy production and consumption systems and electricity companies that supply energy from outside the region. It is possible to assume that there are a variety of energy resources including solar power, wind power, and natural gas as well as energy storage such as that in battery systems, electric vehicles (EV)/plug-in hybrid electric vehicles, (PHEV) and water heaters/coolers within the region.

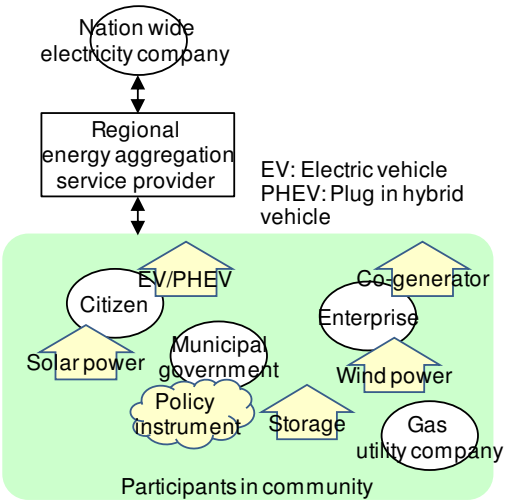


Figure 1. Basic idea of a regional energy aggregation service

The design problem for regional energy aggregation service is defined to clarify functions and to organize stakeholders within the systems to meet the objective of systems.

3 Service Systems Engineering Framework based on Soft Systems Methodologies

In order to realize smarter service systems engineering, engineering methodologies have been developed based on the systems engineering tradition. However these are mainly depending on hard systems engineering for developing mechanical systems even though service systems are definitely composed by social interactions.
Soft systems methodologies have appeared for designing and transforming organizations after the success of hard systems engineering. Soft systems methodology (SSM) proposed by Checkland [3][4] which provided a process and framework for designing and developing systems from social points of views.
As a process model, SSM has the seven stages in which starts by identifying the problematic situation and ends up by taking action to improve the situation. In addition to the process model, SSM provides a structural model for describing an ideal system which is composed of CATOWE standing for the customer, actor, transformation process, world view (Weltanschauung), owner, and environmental constraints.
CATOWE can be seen as elements of service systems where an actor and a transformation process are corresponding to a service provider and a service process respectively. Based on this similarity idea, we can formulate a new service systems engineering framework as depicted by Figure 2 where the seven stages are reduced to 3 steps and CATOWE is structured as a service template.

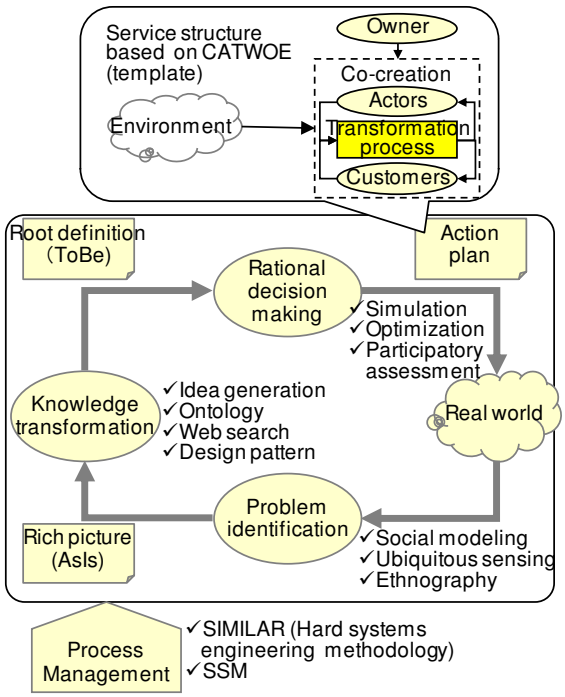


Figure 2. Proposed service systems engineering framework based on soft systems methodologies

4 Design Study for an Aggregation Ser-

vice Systems based on the Proposed Framework

First, the stakeholders of the system were preliminarily identified, who included regional citizens, regional enterprises, gas utility providers, a municipal government, and electricity utility companies who were not within the region in the case of Japan. Second, a rich picture was obtained as the output of the problem identification step that described problems in the management of regional energy services. The results are summarized in Figure 3.

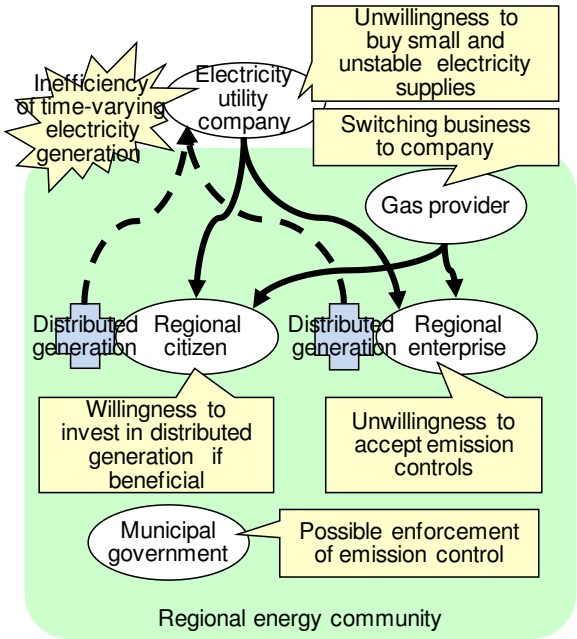


Figure 3. Interests of stakeholders for regional energy aggregation services

(1) Interest of stakeholders

Interests of stakeholders were summarized as follows:

- *Regional citizens*

Regional citizens are willing to save energy and reduce GHG emissions, but are unwilling to pay. However, if they can earn money, they are willing to invest in distributed energy generation and storage.

- *Regional enterprises*

Regional enterprises are willing to invest in energy saving and distributed energy generation if they are beneficial. However, they are unwilling to reduce GHG emissions unless this is enforced by laws.

- *Gas utility providers*

As gas providers are losing citizen customers because of the convenience of electricity, they are switching their businesses to enterprises by appealing to customers about the economy. They expect that laws to enforce reduction of GHG emissions will promote a shift from the use of crude oil to that of gas as the energy resources enterprises.

- *Municipal governments*

Municipal governments presume that they will have to enforce reduced GHG emissions released by regional enterprises in the future, but enforcement is very difficult for the government and also does not imply increased numbers of regional jobs or income.

- *Electricity utility companies*

As electricity utility companies form a nationwide grid in Japan, they are unwilling to buy small and unstable electricity supplies from regionally distributed generators. They argue that buying small and unstable electricity supplies will raise the price of electricity.

(2) Basic idea for the energy aggregation service

Since storage of electricity is quite difficult, facility capacity of electric power supply is determined by maximum demand. The electricity utility companies have been obliged to build their generation and transmittance facilities to meet the maximum demand of which duration is very limited over a year. Since the electricity plants are built to satisfy the peak demand, their operating ratios were very low which result in expensive electricity cost.

Regional energy aggregation service providers could reduce the dynamic components of demand by expecting such collaboration between regional participants. This means that aggregation service providers could separate energy demand into two where the first is flat and is economically supplied by electricity utility companies and the second is variable and is supplied by distributed generation within the region. The CATWOE diagram in Figure 4 is obtained based on this idea.

Basic configuration for an aggregation service provider is shown in Figure 5 based on the CATWOE diagram, where the participants share Ba where energy dealing and transmittance are carried out. Both citizens and enterprises are assumed to have generation as well as storage capabilities that are partially supported by gas providers.

The rich picture and CATWOE diagram have produced a basic structure for a regional energy aggregation service. The importance of problem definition and knowledge integration and emergence to obtain the root definition of the system is visualized in large and complex service systems engineering. It is essential to involve further stakeholders and to integrate diversified knowledge to proceed to the next step.

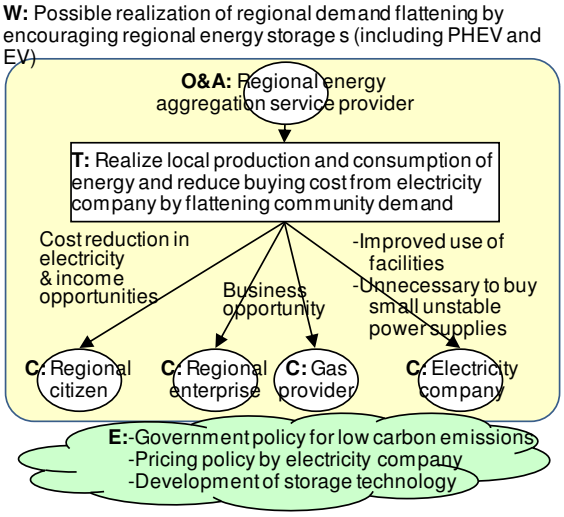


Figure 4. CATWOE diagram for the regional energy aggregation service

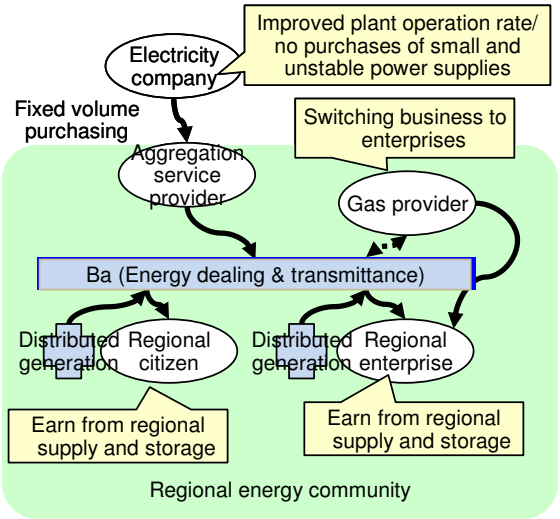


Figure 5. Basic structure of the regional energy aggregation service system

5 Service Co-Creation in Real World Example

We are now having an energy aggregation service in real world. The Tokyo Electric Power Company, Inc. (TEPCO) issued a request in January 2012 for a service to suppress the demand for electricity to be provided to prepare for the high demand for it in summer. The basic idea behind this is outlined in Figure 6. When the service provider called the Aggregator receives a request from TEPCO to suppress demand in this scheme, it needs to reduce the total amount of electricity use by its contracted customers to the requested level.

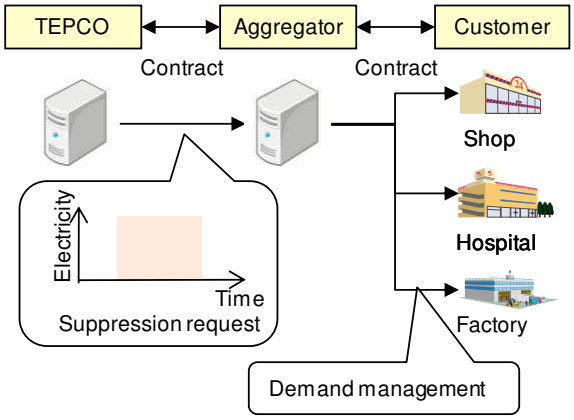


Figure 6. Aggregation service scheme by TEPCO

TEPCO received 81 applications and adopted five applications in March 2012 for this request. The first trial planned to reduce demand by 400 MW in the summer of 2012, and a reduction of 1.4 GW was expected in 2014. Demand management by the applicants differed in five respects as shown in Table 1.

Table 1. Adopted applicants for the demand management services

-Equipping customers with energy management systems
-Introducing a supervisory control system to remotely control customer demands
-Equipping customers with demand monitoring devices and asking customers to suppress electricity consumption
-Equipping customers with private generators that reduce TEPCO's supply on requested occasions
-Introducing integrated management to customer buildings and energy management systems (BEMS) and remotely suppress control

There were a variety of approaches to demand aggregation because it was a first trial and it did not have clear predecessors. The aggregator and customers including TEPCO should innovate their service scheme as is shown in Figure 7 through their experiences. For successful innovation, co-creation of service structures as well as

functions is indispensable. It should also be noted that we are going to have disruptive technologies such as Internet of things [5] which definitely change the structure of the aggregation services.
The service is not static but evolutionary and this fact requires a new approach to service systems engineering. This is a challenge for the next generation of systems engineering. Kosaka et al. [6] have proposed a model of knowledge sharing related to service systems, identification of service fields, knowledge creation for new service ideas, and implementation of service ideas (called KIKI) for this situation. The next section presents new trends toward evolutional systems thinking for understanding the co-creation process.

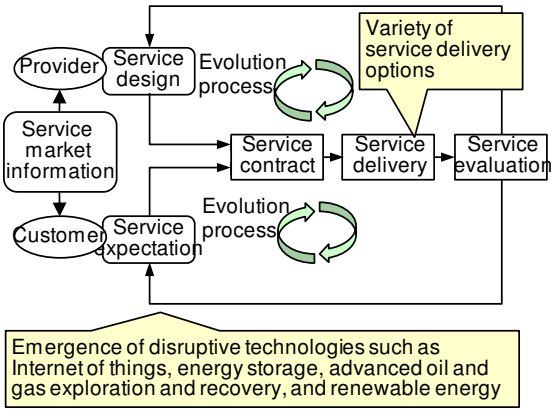


Figure 7. Service co-creation process for the energy aggregation system

6 Transdisciplinary Science and Technology for Preparing Service Co-Creation

Systems approaches including hard and soft systems engineering have significantly contributed to the development of large complex systems and have provided significant potential for reliable service systems engineering. However, we are now facing very difficult situations as our society is so much more tightly linked together. It is quite difficult for us to localize our activities and separate them from the environment, which is changeable because of numerous related unknown factors. We have to be prepared for unpredictable situations, and as our systems should be well planned and flexible enough to endure environmental changes, evolution is an essential property of these systems. Systems engineering is presently challenged by this new situation. We have to tame current acute technological advances such as those in materials, information, and biology that have had significant impacts on our society in addition to satisfying this evolutionary requirement. We need a new foundation for science and technology to cope with unpredictability and make the best use of our advances in scientific knowledge.
Gibbons et al. [7] stated that a new form of knowledge production started from the mid-20th century, which was context-driven, problem-focused, and interdisciplinary. It involved multidisciplinary teams brought together for short periods of time to work on specific problems in the real world. They called this "mode 2" knowledge production.
Kimura [8] claimed that effective multidisciplinary teamwork needs some mechanism and transdisciplinarity is expected to be an essential viewpoint for this mechanism. He and his colleagues claimed that the definition of transdisciplinary science and technology was:
A foundation of science and technology systems that fosters integration of diversified disciplines traversing science, art, and technology under the guiding principle of mathematical and computational logics, and contributes to creating new social values.
Implications of trans-disciplinary science and technology are three-holds as shown in Figure 8 [9]. First it provides overview of diversified heterogeneous knowledge. Second it integrates and organizes heterogeneous elements to a purposeful system. Third it also discovers social wish that gives rational of the purposeful system.

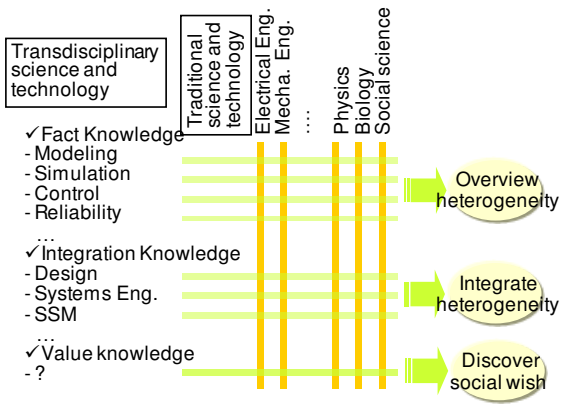


Figure 8. Transdisciplinary science and technology overviewing, integrating, and discovering

In service systems engineering for evolutionary situations, an iterative engineering process in-

cluding problem identification, knowledge transformation, and rational decision making as proposed in this paper should be carried out. Knowledge in transdisciplinary science and technology is expected good resources for the process.

The co-creation process can be seen as a kind of collaborative systems engineering by the customer and the service provider. Development of communication media among the customer and the service provider in terms of the systems engineering is essential for successful design of the co-creation process.

7 Concluding Remarks

Systems engineering has significantly contributed to developing large and complex artifacts. Service systems are providing new opportunities for the systems engineering. The present paper has proposed a new service systems engineering framework based on soft systems methodology.

The proposed framework has been applied to developing a regional energy aggregation service system and has been confirmed its effectiveness. Future of the energy aggregation service system has been discussed and importance of arranging a co-creation process has been pointed out. For the successful arrangement of co-creation process, value of transdisciplinary science and technology has been presented which is emerging in 21st century.

Acknowledgment

This research is partially supported by Japanese Grants-in-Aid for Scientific Research (KAKENHI No. 25240049).

References

[1] H. Chesbrough & J. Spohere, A Research Manifesto for Service Science, *Communications of the ACM*, 49, 7, 35-42, 2006

[2] S. L. Vargo & R. F. Lusch, Evolving to a New Dominant Logic for Marketing, *Journal of Marketing*, 68, 1, 1-17, 2004

[3] P. B. Checkland, *Systems thinking, systems practice*, John Wiley & Sons, 1981

[4] P. B. Checkland & J. Scholes, *Soft Systems Methodology in Action*, John Wiley & Sons, 1990

[5] Mckinsey Global Institute, *Disruptive technologies: Advances that will transform life, business, and the global economy*, http://www.mckinsey.com/insights/mgi, 2013

[6] M. Kosaka, Q. Zhang, W. Dong & J. Wang, Service value co-creation model considering experience based on service field concept, *Proc. of 2012 9th International Conference on Service Systems and Service Management (ICSSSM)*, 724-729, 2012

[7] M. Gibbons et al., *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, Sage Publications of London, 1994

[8] H. Kimura, What is Transdisciplinary Science and Technology?, *Oukan (Journal of Transdisciplinary Federation of Science and Technology)*, 1, 1, 4-8, 2007 (in Japanese)

[9] M. Funabashi et al., The 10 Year History of Transdisciplinary Federation of Science and Technology– From Concept Making to Practices –, *Oukan (Journal of Transdisciplinary Federation of Science and Technology)*, 7, 1, 24-35, 2013 (in Japanese)

An Enhancement of the Stakeholder Modeling Scheme for Requirements Engineering of Software Systems for Public

Jidtima Sunkhamani

Software Engineering Program, School of Informatics
Walailak University, Thailand
sjidtima@wu.ac.th

Abstract

Software systems for public are the systems initiated mainly for the benefits of a wide range of people. Such a system is initiated and its application domain and context should be focused early even before proposing it to development. Stakeholders are the main sources of software requirements and also domain knowledge. In order to include stakeholders in the requirements process judiciously, they have to be identified together with the reasons for their inclusion. This paper reports on an adaptation of Stakeholder Modeling Scheme to enhance it for the requirements engineering of software systems for public. The adapted scheme was applied to a case study, Krung-Shing Climate Monitoring system, a software system initiated for the benefits of the public. The results showed that the adapted scheme helped in stakeholder identification and requirements engineering of such a system.

Keywords: Stakeholder Modeling Scheme, Stakeholder Diagrams, Requirements Engineering, Software systems for public, Application domain study, Krung-Shing Climate Monitoring

1 Introduction

System stakeholders are parts of a system and its environment. They are the main source of domain knowledge and system requirements; thus, achieving good quality of requirements is impossible without involving stakeholders [1]. Many methods and techniques have been proposed to help requirements engineers in requirements elicitation. Applying the many existing methods and techniques does not warrant that the engineers will obtain requirements, which reflect the "real" needs unless the "right" stakeholders are identified and involved. However, including all or too many stakeholders is impractical because a project is always constrained by schedule and budget [2]. Hence, it is necessary that the engineers know why some specific stakeholders are more relevant or important to the system than the others. Without insightful information, the engineers may elicit requirements from the wrong stakeholders, approach the right ones but fail to elicit specific and clear requirements from each of them, and/or interpret what each stakeholder says or implies incorrectly.

There are cases that some systems are initiated, proposed, or recommended because of their potential for solving some specific problems of or bringing some opportunity to the public. Often they are technology- or market-driven systems. They may be initiated by engineers' creativity or by someone's interest in standard improvement of society. We then call such systems "software systems for public".

The scope of a system for public can be unclear at the outset but it tends to be beneficial to a wide range of stakeholders rather than a particular organization. Hence, its stakeholders are likely to cross many organizations' boundaries and may cover people in public. Before proposing such a system to a project sponsors, the initiator should study its application domain. Thus, its stakeholders must be identified as early as possible.

To date, many methods for stakeholder identification have been introduced. Sharp et al. suggested the engineers to explore the supplier, client, and network of 4 predefined classes of stakeholders: users, developers, legislators, and decision makers [3]. This method identifies stakeholders from the outset, so it works for systems having predefined and clear boundary. Gause and Lawrence proposed to use brainstorming to identify stakeholders [4]. Stakenet adopted Sharp's method to acquire the initial set

of stakeholders' roles and to have the identified stakeholders recommend other stakeholders similar to snowball technique [5]. Onion model defines a classification of stakeholders [6]. The method provides predefined stakeholder classes with no means to find who the real stakeholders are. When they are referred to as 'users' or 'customers' of a system, it is not clear who they really are in the subject domain of the system under study. All [3], [4], [5], and [6] were designed for studying the systems to serve the organizations that own them. Ballejos and Montagna presented an approach to stakeholder identification in inter-organizational environment. It first distinguished stakeholders to those internal and external to the firm and sub-classified them using 4 criteria, namely, function, geography, hierarchy, and knowledge before associated them with the roles they play [7]. Stakeholder roles were not provided in the method but general classes such as those defined in Onion model were suggested. Similarly, this approach was designed for the systems that have clear boundary and host organizations. Furthermore, all of these methods identify stakeholders without the information about why each stakeholder should be included in the process.

Sunkhamani [8] proposed Stakeholder Modeling Scheme for software systems engineering. The scheme also suggests that the proposed system belongs to a "focal organization", the organization that owns the system as it hosts the development of the system. Nonetheless, unlike others, it offers a way to identify stakeholders alongside the reasons why they are the stakeholders of the system. It contains Stakeholder Modeling Language, enriched with the model elements enhancing the engineers to capture application domain knowledge while eliciting the requirements of stakeholders. However, the fact that the language comprises a rich set of model elements makes the modeling process time-consuming and the resulting model complex. To reduce the cost of applications for the study of a system at the early stage, a more compact process is needed.

This paper presents the enhancement of the Stakeholder Modeling Scheme for requirements engineering of software systems for public. It contains 6 sections. The Stakeholder Modeling Scheme is presented in Section 2 as background. Section 3 describes the adapted modeling language and process. Section 4 presents the result of the application to a case study. Discussion is given in Section 5, followed by conclusion in Section 6.

2 Background

2.1 Stakeholder Modeling Scheme

A stakeholder is "an organizational body, an individual, or a group of people who depend on the system, can be affected by the system, or who influence system requirements" [8].

Stakeholders are categorized by their concerns into three types: domain stakeholders, project/process stakeholders, and product stakeholders. Their concerns are in the subject domain, the development project/process, and the use of the system, respectively.

Stakeholders are modeled with Stakeholder Modeling Language. The model elements in the language are entities, relationships between entities, and attributes. Entities (or concepts) are grouped into core and supporting entities. The core entities are Stakeholder, Organization, FocalOrganization, Individual, Role, Task, and Resource. The two supporting entities added to the language to help in the modeling process are CollaborationSetting and Layer. Relationships link entities. Attributes capture the relevant details of entities and relationships. Most, if not all, of the language elements are included in the adapted scheme and their semantics are described in the next section.

The modeling elements can be represented in textual or graphical form. The textual model can present more information while the graphical model is more compact. As trade off for simplicity, the graphical form contains a subset of the language elements. In most cases, the information captured graphically is enough to help the engineers to understand the problem domain. Particularly for the early stage of the domain study, the graphical form is suitable. We adapted it in this work.

2.2 Stakeholder Diagrams

The graphical stakeholder model is presented as Stakeholder Diagrams. There are three different kinds, namely, Collaboration Structure Diagram (CSD), Task Exploration Diagram (TED), and Stakeholders Mapping Diagram (SMD).

CSD presents the structure of the collabora-

tions in tasks of all stakeholders, called the "collaboration structure". It shows all collaboration settings and tasks included in each one, all tasks only at the highest level of detail, relevant stakeholders, and resources owned by each stakeholder.

TED presents the concepts related to the tasks displayed in CSD. It is the exploded version of CSD in a way that it presents tasks at all levels of detail including resources that each task uses and produces. However, as resources owned by each stakeholder are shown in CSD, they are not duplicated in TED.

SMD captures all revealed stakeholders and relationships between them. The number of stakeholders modeled in SMD may exceed that modeled in CSD because SMD contains all disclosed stakeholders, although they are not related to the collaboration structure.

For separation of concerns, each kind of diagrams can be modeled in three different worlds: subject, development, and usage worlds, focusing on elements in the problem domain, the development project, and the system in operation, respectively. Each type of stakeholders, namely, domain, project/process, and product stakeholder type is the subject to be modeled in each of the three worlds respectively. Moreover, the subject world is separated into two different sub-views, namely, view of the "system as is" and the view of "system to be". So CSD can be drawn in 4 diagrams, i.e., CSD-S-AsIs, CSD-S-ToBe, CSD-D, and CSD-U to capture the collaboration structures in the subject world as it is, the subject world as it will become when the new system is in place, the development world, and the usage world, respectively. As well as CSD, TED and SMD each can be drawn in 4 different views.

3 The Adapted Stakeholder Modeling Scheme

This section describes the contribution of this work to make Stakeholder Modeling Scheme applicable for requirements engineering of software systems for public.

3.1 Modeling Elements in Stakeholder Diagrams

The entities, relationships, and attributes elements can be modeled in the stakeholders diagrams are described as followed. (For understandability, presenting each of them for the first time, we boldface the **Entity names** (and as Entity the following times), underlined the <u>relationship names</u>, and italic the *attribute names*. In addition, for clarity, the value of attribute in the language explanation is presented in single quotation marks whereas the value of some entities and attributes in examples are presented in double quotation marks.)

The entity at the center of the language is **Stakeholder**. It is classified into 3 subclasses, namely **Role**, **Individual**, and **Organization**. Stakeholder is an abstract type entity, meaning the type of each Stakeholder instance must be defined when modeled. The Role type Stakeholder refers to a set of stakeholders who have the same responsibility for behavioral characteristics, for example, "Student". Individual is a type of Stakeholder referring to individual persons, "Daniel", for example. The Organization type stakeholder can be any organization related to the system under study.

A Stakeholder of any type may <u>play</u> a Role type stakeholder. The stakeholder playing the role is called *rolePlayer*. For example, "Daniel", an Individual type stakeholder, is a rolePlayer of "Student", a Role type stakeholder. Each Role has the *playerType* indicating the number of the players of the role. Depending on the number of the players of the Role, the value of playerType can be 'only one', 'one', 'many', or 'group'. They mean the role is played by one and only one, one, many, or a group of players, respectively. For the only one and one player Type, the name of the stakeholder who plays the Role must be specified. For the many and group playerType Roles, some, if not all, players should be specified unless the engineer knows that it is not useful to model them.

An Organization type stakeholder has some Stakeholder(s) as its *representative.*

Stakeholders may relate between themselves in 4 kinds of relationships, namely <u>reports</u>, <u>consults</u>, and <u>liaises</u> giving 3 attributes with the same names, i.e., *reports*, *consults*, and *liaises*. The reports relationship implies direct influencing power of the superior Stakeholder over the subordinate one. The consults relationship covers some degree of influencing power between them although less formal or informal. The engineers model these relationships when they see some influencing power between them from any spe-

cific circumstance. Liaises is the relationship which shows an informal communication link between Stakeholders where there is no influencing power found between them.

A **Task** represents an association between two groups of Stakeholders, namely the task *owner* and the task *collaborator*. The owner has the ultimate responsibility for completing the task whereas the collaborator participates and has responsibility for their participation. Many stakeholders may share responsibility in a Task. Some Task may not have a collaborator if the owner can complete it by oneself, a so-called 'self- completing task'. If the owner and the collaborator of a task are specified with the same role, it means that different role players of the role collaborate in the task. A Task may have a *regulator* who defines a particular way to do the task.

A Task *uses* or *produces* some resources. A Task may be decomposed into smaller Tasks. In other word, a Task *includes* other Tasks.

Resources can be a system or some kind of information in any format. A Resource may *include* other Resources.

The owner of a Resource (called *rOwner*) is the Stakeholder who owns it. The privacy owner (called *pOwner*) is the Stakeholder whose privacy data are captured in the Resource. Some Resources do not contain privacy data, in which case, it has no privacy owner. For the system type Resource, its *maintainer* is the Stakeholder who maintains it.

The graphical notations of the language elements are presented in Figure 1.

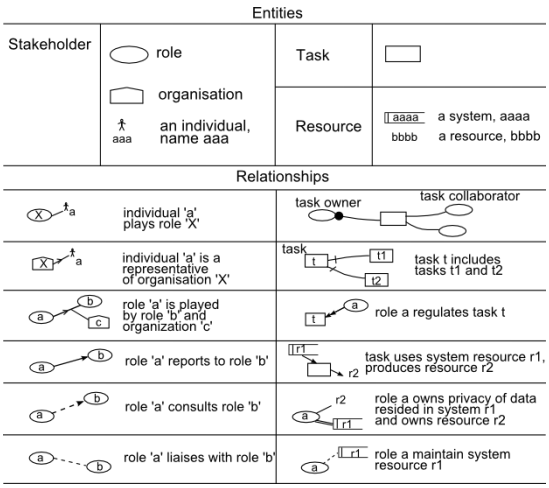


Figure 1. The graphical notation of the elements

3.2 The Adapted Modeling Process

As the intention of the initiator of a system for public is to solve some problems for or give some opportunities to the public, their main focus of study is at the situation in the subject domain. The engineer should identify domain stakeholders, with no concern at project/ process/product stakeholders, to capture as much application domain knowledge as possible. Moreover, to reduce time and complexity, the as-is state and the to-be state of the subject world are not separated into different diagrams. Therefore, there are three diagrams to be modeled: CSD-S, TED-S, and SMD-S.

The modeling process is iterative so that more stakeholders are uncovered in each revision. Engineers model stakeholders concurrently with eliciting their requirements. The modeling process is as followed:

Step 1: From the problem statement, model outstanding domain stakeholders in SMD-S. In any cases, it should not be difficult to identify the very first stakeholders. They normally are obvious individuals or roles exist in the subject domain.

- For each individual, identify the roles they play, which are relevant to the problem.
- For each identified stakeholders, identify the tasks they relate to, i.e., the tasks each one owns, collaborates, and/or regulates. Model the tasks in TED-S and CSD-S. The tasks to identify are those relevant to the problem at hand. The statement of the problem is helpful for revealing them.

Step 2: Identify the collaboration setting(s) relevant to the problem domain. Model them in CSD-S.

- For each collaboration setting, reveal the tasks of which it is composed. Modify CSD-S accordingly. Moreover, all the revealed tasks must be modeled in TED-S.
- For each identified task, identify the task owner, collaborator, and regulator, (if known and/or exist). In addition, identify the resources used and/or produced by each task. Modify TED-S accordingly.
- For each revealed resource, identify stakeholders who are the owner, privacy owner, and maintainer (if exist) of the

resource in CSD-S. Model the resource used and produced by the respective tasks in TED-S.

- For each revealed stakeholder, add it in SMD-S. Depending on the type of the stakeholder, model the roles the individual type stakeholder plays or the player(s) of the role type stakeholder or the representative of the organizational type stakeholder. Identify and model relevant relationships between stakeholders in SMD-S.

Step 3: Identify the collaboration setting(s) relevant to the goal of the proposed system. Repeat *Step 2* until all the elements in the collaboration setting are uncovered.

Step 4: Gathering domain knowledge and stakeholders' requirements. Revise the diagrams by repeating *Step 2* until enough domain knowledge is gained and all important requirements for the system are clear.

4 Results: An Application to a Case Study

4.1 The Case Study

The adapted modeling scheme was applied to a software system for public, Krung-Shing Climate Monitoring system. The problem statement was that Krung-Shing, a sub-district in Nakhon Si Thammarat, had not enough information to enhance the inhabitants to deal with a natural disaster, particularly heavy rainfall causing severe flood and landslides, in time for them to safely evacuate from the villages. From the last incident in 2009 when many areas were badly destroyed by flood and landslides, the inhabitants could not rely only on the authorities. During the disaster most, if not all, communication channels were disconnected. To better prepare for such incidents in the future, the system was initiated by a person living in the area.

The goal of the system was to provide basic and geographical information of the villages in Krung-Shing and information about the climate that helps the inhabitants and also the authorities to predict the likelihood of severe flood and landslides during a period of heavy rainfall. The system would be used mainly for monitoring rainfalls and supporting the evacuations if needed. The ultimate goal was to decrease loss of lives and properties from possible future incidents.

4.2 Data Collection and Modeling Process

The data were collected by interviews. The first stakeholder identified was the initiator of the system, Mr.Rangsan. From the first few pieces of data collected, the first version of SMD-S was drawn as illustrated in Figure 2.

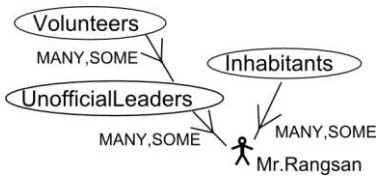


Figure 2. The first version of SMD-S

The diagram presents that Mr.Rangsan was one of the "Inhabitants" and also "UnofficialLeaders". During the disaster, many unofficial leaders played the "Volunteers" role. The keyword "MANY" means the role can be played by many players; whereas "SOME" means some of the players are presented in the model. One of them was Mr.Rangsan.

The first collaboration setting, "Evacuate", was modeled in CSD-S. The stakeholder model led us asked the questions regarding the tasks, resources, and stakeholders existed in the collaboration setting. We found that it contained one task, i.e., to evacuate the inhabitants. Thus, the TED-S was appeared to be the same as CSD-S, as presented in Figure 3.

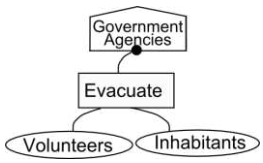


Figure 3. The first version of CSD-S and TED-S

In the figure, Evacuate is a task responsible by some government agencies, in collaboration with inhabitants and volunteers. At the same time, new pieces of data were added to the SMD-S, as shown in Figure 4.

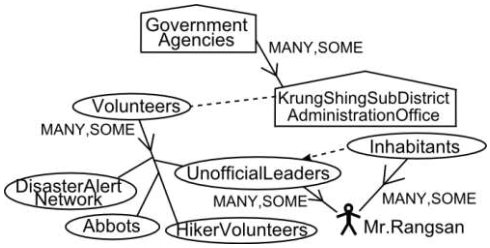


Figure 4. The second version of SMD-S

The diagrams modeled so far partially described the subject domain as it was. Then the "to-be" state of the subject domain was analyzed. Together with the goal of the system, a collaboration setting "Monitoring and Evacuation", a collaboration structure for the preparation of evacuation, alert and evacuation was identified. We then added the "Monitor" task into CSD-S. Mr.Rangsan added that, in doing so it was necessary to have information about rainfall and other basic data about the villages such as geographical aspects of the villages, data of inhabitants and households, pre-defined evacuation points. So we continued by asking about the sources of the data he mentioned. The information collected were used for revising CSC-S, as shown in Figure 5.

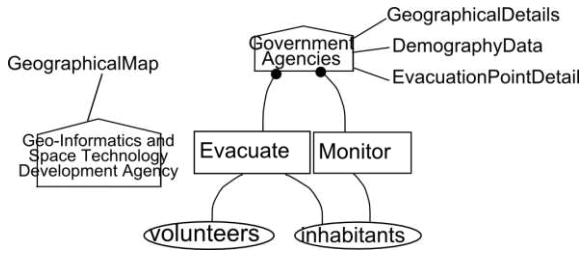


Figure 5. The second version of CSD-S

The data required for carrying out each task were identified and TED-S was updated, as depicted in Figure 6.

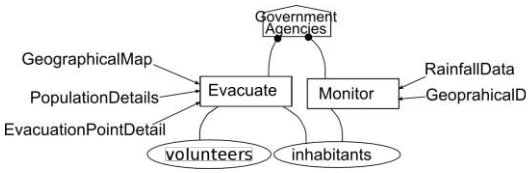


Figure 6. The second version of TED-S

More stakeholders identified during the modeling process were interviewed. Data were collected from Krung-Shing Community Head, a Community Development Officer, a Disaster Prevention and Mitigation Officer, and a Police Assistant Volunteer. The model was used as the guideline of the interview questions. The stakeholder diagrams of the subject world, namely, CSD-S, TED-S, and SMD-S were updated after each interview to capture new domain knowledge. The results after some more rounds are presented in Figure 7, Figure 8, and Figure 9, respectively.

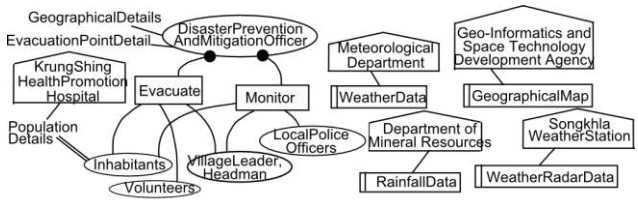


Figure 7. A refined version of CSD-S

Figure 8. The refined version of TED-S

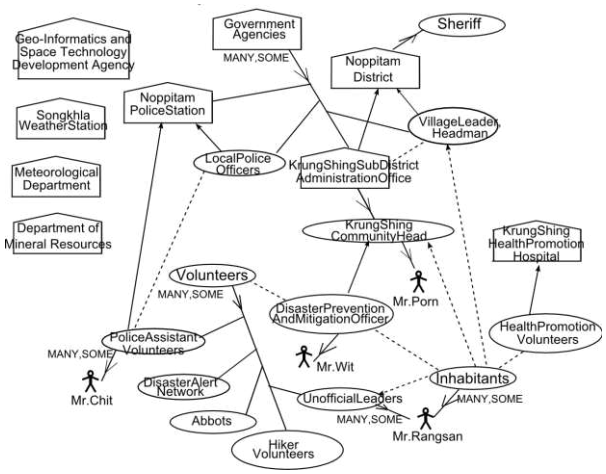


Figure 9. A refined version of SMD-S

While collecting the domain knowledge, the requirements of those stakeholders were elicited. Some examples are presented in table 1.

5 Discussion

During the course of study, many potential stakeholders were revealed and some of them were met for interviews. The stakeholders and their relationships are shown in SMD-S. The relationships between stakeholders and tasks including resources relevant to the problem are captured in CSD-S and TED-S. Applying the adapted modeling scheme to the case study, we found that the modified scheme had these advantages:

- The adapted model helped the engineer to focus the stakeholders in the problem domain, hence, it is suitable for studying a system at the very early stage, as this one for public. The separation of stakeholders into three different views in [8] helped in identifying domain stakeholders with no concern or question about who would be the project/process/product stakeholders.
- Developing the model of the subject world "to-be" from the subject world "as-is", i.e., drawing different versions of one diagram, decreased errors, time, and complexity should have incurred from drawing many versions of two different diagrams. At the end, the view of the world when the system is in place is what the engineer must have. The engineer simply developed the changing view of the world from the current state through the process. For example, from Figure. 3 to Figure. 5 and to Figure. 7.
- The modified scheme was useful for studying the application domain of the case study because it could be applied without identifying the host organization or knowing the scope of the system. This means that engineers can get benefit from modeling the application domain of software systems for public even before the start of a development project.
- The adapted modeling process was done manually. This means that it was systematic and simple enough to use by a skilled modeler. In addition, the resulting diagrams are meaningful. However, a new adopter has to learn to be familiar with the modeling language. In any case, for the method to be adoptable in a more complex subject domain, a modeling tool is needed to help the process practical, less time consuming and less error prone.

The adapted scheme still offers its benefits to engineers, i.e., to systematically collect domain knowledge, no lesser than the full scheme does. At the same time, with the information suggested in the model, requirements can be elicited from identified stakeholders effectively. For example, with the essential information about their relationships to other stakeholders and other entities, such as Tasks and Resources from the model, we could ask more relevant questions, expect the scope of the answers, and so interpret requirements stated by the stakeholders more meaningfully. Furthermore, the diagrams gave the big picture of the application domain, which was helpful for making decision about the scope of the system. For example, in Evacuate task, if the geographical map of the sub-district was really needed, Geo-Informatics and Space Technology Development Agency, the owner of the map data, must have been consulted as a key stakeholder. Otherwise, the scope of the system must have been decreased or other options had to be found.

Table 1. Examples of stakeholders identified during the modeling process with some of their requirements

Stakeholders	**Domain knowledge gained from the model**	**Stakeholder's requirements**
Inhabitants	• They collaborate in the evacuation and monitoring tasks. • Their privacy data are kept in the populationDetails. • They usually respect the official and unofficial leaders. • They are closed to the health promotion volunteers.	The system shall.. • alert the inhabitants in advance when the situation is becoming serious. • contain knowledge about geography, demography, evacuation process, including knowledge helpful for them to predict severe flood or landslide, e.g., quantity of rainfall, cloud observation, etc. • provide information about the persons who need special care during the incident such as elderly.
Unoffical Leaders	• They are a volunteer group. • They gather to help in evacuation during the incident.	The system shall.. • provide geographical map with details enough for planning the evacuations when needed. • give precise details of the inhabitants in needed of special help.
Disaster Prevention and Mitigation	• They are responsible for the monitoring	The system shall.. • automatically warn the inhabitants when

Stakeholders	Domain knowledge gained from the model	Stakeholder's requirements
Officers	and evacuation tasks. • They report to Krung-Shing Community Head.	the situation is risky. • record loss information and send it directly to the office of the district.

6 Conclusion

Being the main source of domain knowledge and requirements, stakeholders must be correctly identified and properly involved. Stakeholder Modeling Scheme offers stakeholder identification alongside insightful information regarding each stakeholder which is very beneficial to the engineers because it shows the degree of relevance that each stakeholder has to the system under consideration. The model is both descriptive and prescriptive (i.e., it captured relevant details of stakeholders explored during the process and led to identifying them). These benefits are not offered in other approaches to stakeholder identification. However, the scheme is not applicable for requirements engineering of systems for public. Such a system has some special characteristics. It is initiated because of its potential to be useful to a wide range of stakeholders. The scope of such a system has not been decided and it tends to cross many organizational boundaries including people in public. Its application domain should be understood even before proposing to a development project sponsor, namely, before knowing the host organization.

We have presented the adaptation of Stakeholder Modeling Scheme to make it applicable for requirements engineering of systems for public. The adapted scheme applied to an initiative, Krung-Shing Climate Monitoring system, showed that it offered a systematic way to identify stakeholders alongside their relationships to other stakeholders and to other elements in the system and the model suggested more questions and requirements to ask each of them. In addition, the adapted version is more compact and can be applied before identifying the host organization. Hence, the adapted scheme is applicable for requirements engineering of systems for public.

Acknowledgment

We would like to thank Walailak University for the research grant and the participants (who do not want to have their names presented) for their contribution as the sources of data for this study.

References

[1] Beyer, H. & Holtzblatt, K, Contextual design, *Interactions* 6(1), 32–42, 1999.

[2] Gause, D. & Weinberg, G., Exploring Requirements: Quality Before Design, Dorset House Publishing Co., Inc., NY, USA., 1989.

[3] Sharp, H., Finkelstein, A., Galal, G., Stakeholder identification in the Requirements Engineering process, *DEXA Workshop*, 1999.

[4] Gause, D. and Lawrence, B., User Driven Design, *Software Testing & Quality Engineering*. 1(1), 22–28, 1999.

[5] Lim, S., Quercia, D., Finkelstein, A., StakeNet: Using social networks to analyse the stakeholders of large-scale software projects, *in Proceeding of the ACM/IEEE 32nd International Conference on Software Engineering*, Cape Town, South Africa. ACM Press, 2000.

[6] Alexander, I. and Robertson, S., Understanding project sociology by modeling stakeholders. *IEEE Software.* 21(1), 23–27, 2004.

[7] Ballejos, L. and Montagna, J., Method for stakeholder identification in interorganizational environments. *Requirments Engineering.* 13(4), 281–297, 2008.

[8] Sunkhamani, J. *Modelling stakeholders in software systems engineering.* Ph.D. Thesis. University of London. UK, 2010.

Learning and recognizing behavior patterns based on the distribution of the direction of movement

Kizuku Nishimura[†] **Shigeki Aoki**[‡] **Takao Miyamoto**[‡]
[†]Graduate School of Engineering, Osaka Prefecture University
[‡]College of Sustainable System Sciences, Osaka Prefecture University

Abstract

Automatic video surveillance systems have recently become an active area of research. Typical methods of research can be divided into two categories: approaches that provide model data for abnormal behavior, and approaches that detect rare behavior. In the former approach, methods can detect the behavior given by the model data, but, similar to the latter approach, they cannot detect all types of abnormal behavior. In this paper, we propose a novel method using approaches that detect rare behavior. First, we extract comprehensive movement flows of people from a surveillance camera video. Second, we classify the behavior patterns into a number of clusters. Finally, when a new flow is observed, we identify whether it is "normal" or "irregular" by referring it to the defined clusters.

Keywords: Behavior pattern recognition, Detection of irregular situations, Comprehensive flows, Automatic video surveillance

1 Introduction

In recent years, a large number of surveillance cameras have been installed to prevent crime and monitor traffic. In order to confirm whether an observed situation is abnormal, supervisors constantly monitor the video feed from the cameras. However, an abnormal situation may be missed because each supervisor must monitor the video from many cameras. There are two typical methods for solving this problem. One is to reduce the number of videos each supervisor must monitor, and the other is to automatically detect abnormal behavior. Stauffer et al. [1] reported a method for analyzing and displaying the behavior of people and traffic using surveillance cameras with an overlapping field of vision, and Lee et al. [2] proposed a technique for recognizing and displaying traffic by tracking vehicles between multiple surveillance cameras. These methods can reduce the number of videos that must be monitored by the supervisors. Research into the detection of abnormal behavior can be divided into two categories. One is the detection of abnormal behavior based on a set of given abnormal behavior models, and the other is the detection of rare behavior. Saxena et al. [3] developed a typical technique for the detection of abnormal behavior. By applying a Scenario Recognition Engine model to the trajectories of people, they were able to detect behavior defined as abnormal, but could not classify all abnormal patterns. Mahadevan et al. [4] proposed a method for detecting the anomaly behavior by using a model is based on mixtures of dynamic textures. Gibo et al. [5] have proposed a method for detecting irregular situations based on comprehensive flows. Unfortunately, if we use these methods, the behavior patterns must be relearned when the number of surveillance cameras is increased or their position is changed a little.

In this paper, we propose a method that detects irregular situations by learning behavior patterns. A comprehensive set of behavior patterns is first extracted, and these are then learned by classifying them. Finally, we detect irregular situations by referring to the defined clusters. Because our method uses features that are independent of location-based information, the existing problem of relearning when the position of the surveillance cameras changes is resolved. Thus, we believe that our method is applicable to a wide range of environments.

2 Extraction of comprehensive flows

2.1 Extraction of comprehensive movements

It is difficult to correctly extract the movement flow of an individual object in a crowded place, because frequent occlusions and noises occur. Thus, we do not extract individual movements, but rather the comprehensive movement. Extracting comprehensive movement has the effect

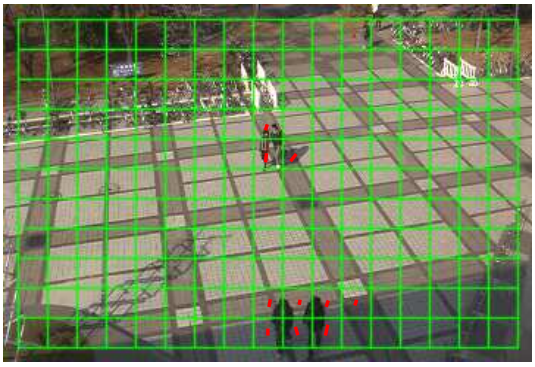


Figure 1. Example of comprehensive behavior

of reducing noise and aiding the detection of abnormal flows, because the individual movements affect the comprehensive movement.

Firstly, to extract the comprehensive behavior, we calculate the optical flow of each pixel in the input image. Secondly, we divide the image into 20×20 [pixel] segments. This segment size is chosen because it can capture the motion of a single person as an optical flow. Finally, in each segment, we calculate the mode values of the optical flows. Figure 1 shows an example in which the movement depicted by the arrows in the center of each segment is extracted. The movements $\boldsymbol{v}_n = \{\boldsymbol{u}_n, \boldsymbol{v}_n\}(n = 1, 2, \ldots, N)$ are extracted by carrying out the above procedure, where N is the number of segments. A comprehensive movement $\boldsymbol{V}_t = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ is created by raster-scanning $\boldsymbol{v}_n(n = 1, 2, \ldots, N)$.

2.2 Division of comprehensive flows [5]

In this study, the following comprehensive flows are observed in crowded places.

I. People appear in sight of a camera and then disappear out of sight

II. The degree of congestion changes

III. The direction of people's movement changes

We regard the moment at which people appear in sight of the surveillance camera, the moment when the degree of congestion changes, and the moment when the direction of movement changes as the start of a new comprehensive flow, and we call this a division point. We set the division points in the sequence of comprehensive movements

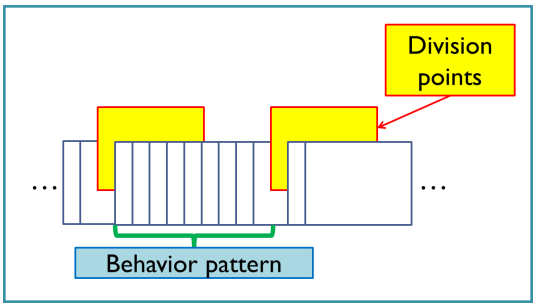


Figure 2. Example of a behavior pattern

$\{\ldots, \boldsymbol{V}_{16}, \boldsymbol{V}_{17}, \ldots, \boldsymbol{V}_{21}, \boldsymbol{V}_{22}, \ldots\}$, and define the comprehensive movement between the division points as a comprehensive flow. We denote the $p-$th extracted comprehensive flow as $\boldsymbol{I}_p = \{\boldsymbol{V}_1^p, \boldsymbol{V}_2^p, \ldots, \boldsymbol{V}_{T_p}^p\}$, where T_p is the length of the sequence $\boldsymbol{I}_p$. The following describes the determination of a division point. In this method, as shown in Figure 2, we define the images between the division points as a behavior pattern.

· **Division Point Determination by Appearance** When $\boldsymbol{V}_t = \boldsymbol{0}$ is observed for the fourth time in a row before becoming non-zero, we set a division point for the start of a comprehensive flow. In addition, when $\boldsymbol{V}_t = \boldsymbol{0}$ is observed for the fourth time in a row, we set a division point for the end of a behavior pattern.

· **Division Point Determination by Degree of Congestion Change** First, we calculate the number of non-zero values $\boldsymbol{v}_n$ in $\boldsymbol{V}_t$, where $\boldsymbol{v}_n \neq \boldsymbol{0}$ denotes that people exist in segment n. The degree of congestion h_t is calculated using the sum of $\boldsymbol{v}_n \neq \boldsymbol{0}$:

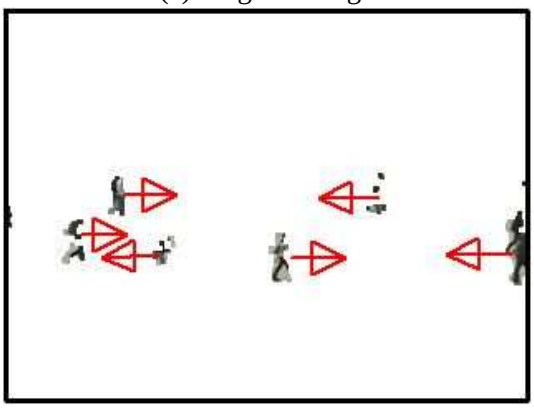
$$h_t = \sum_{n=1}^{N} \begin{cases} 1, & \text{if } \boldsymbol{v}_n \neq 0 \\ 0, & \text{otherwise} \end{cases}. \tag{1}$$

Second, when h_t meets the following requirement, we consider the time t to be a division point within a sequence of feature vectors, where α is a threshold value:

$$\log(|h_t - \frac{1}{5}\sum_{m=1}^{5} h_{t-m}|) > \alpha. \tag{2}$$

In the above equation, to prevent a false division point due to the incorrect extraction of movements, we calculate the absolute difference between the average of $h_{t-1} \sim h_{t-5}$ and h_t. Five

(a) Original image



(b) Extracted foreground and optical flow

Figure 3. Estimation of foreground

frames are used because an exploratory experiment found that this was the optimal number.

· **Division Point Determination by Changes in People's Direction** In order to survey the similarity between people's direction at time $t(\boldsymbol{V}_t)$ and at time $t-1(\boldsymbol{V}_{t-1})$, we use Wilks' λ, which is a kind of analysis of variance, and M-BOX, which analyzes the average value with a significance level of 5%. The null hypothesis is that the "distribution of $\boldsymbol{V}_t$ and $\boldsymbol{V}_{t-1}$ is identical," and the alternative hypothesis is that the "distribution of $\boldsymbol{V}_t$ and $\boldsymbol{V}_{t-1}$ is different." When the null hypothesis is rejected, the alternative hypothesis is adopted and we set a division point.

3 Learning and recognizing behavior patterns and the detection of irregular situations

We now discuss how to learn and recognize the behavior patterns extracted in Section 2.

First, we extract feature vectors based on the distribution of the direction of movement from each behavior pattern. Second, we apply a weight to each element of this feature vector. We regard these feature vectors as coordinate values of the behavior patterns, and classify them using the Mean-Shift technique. When a new flow is observed, we then detect irregular situations with reference to the defined clusters.

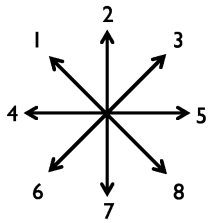


Figure 4. Directions of quantized comprehensive flows

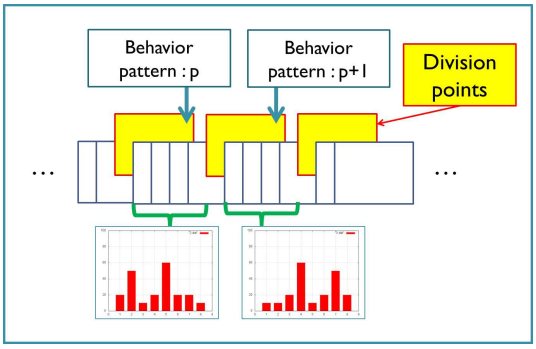


Figure 5. Feature vector based on the direction of movement

3.1 Learning of behavior patterns

In order to extract the comprehensive movement of people, each image in a behavior pattern is estimated from its background by a Gaussian mixture model. As shown in Figure 3, we extract the foreground and calculate the optical flow in each pixels. We then extract the following two feature vectors from the optical flow.

· **Feature vector based on the direction of movement** First, we divide the image into 20×20 [pixel] segments and quantize the flow into eight directions, as shown in Figure 4. Second, we calculate the mode value of the direction in each segment, and process all images in each behavior pattern. As shown in Figure 5, we then calculate the frequency of each direction, before normalizing and sorting in descending order (Figure 6).

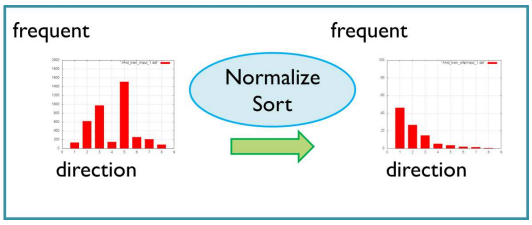


Figure 6. Normalize and sort

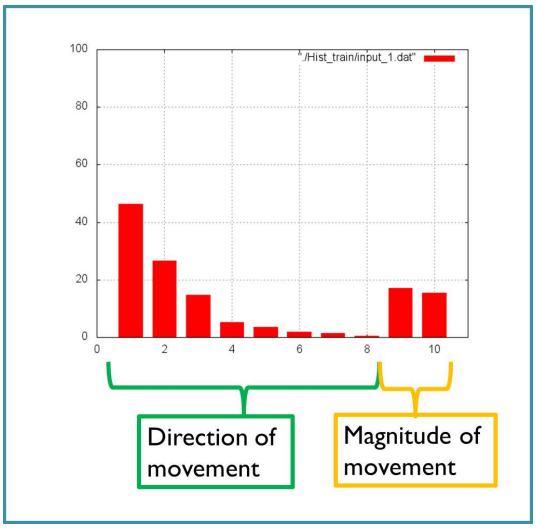


Figure 7. Ten-dimensional feature vector

· **Feature vector based on the magnitude of movement** In order to consider the speed of movement, we calculate the average and variance of the norm of the optical flow in all images for each behavior pattern.

We assign a weight to each element of the two feature vectors obtained above. As shown in Figure 7, we obtain a 10-dimensional feature vector in which the weight represents the direction and speed of movement. The first feature vector is large when the frequency of movement is large. The second feature vector clarifies the magnitude of the movement. We regard these 10-dimensional vectors as the coordinate values of the behavior patterns, and classify the behavior patterns in 10-dimensional space using a Mean-Shift with a Gaussian kernel. The number of clusters is determined by Mean-Shift.

3.2 Detection of irregular situations

First, when a new behavior pattern is observed, we calculate its coordinate value by the method

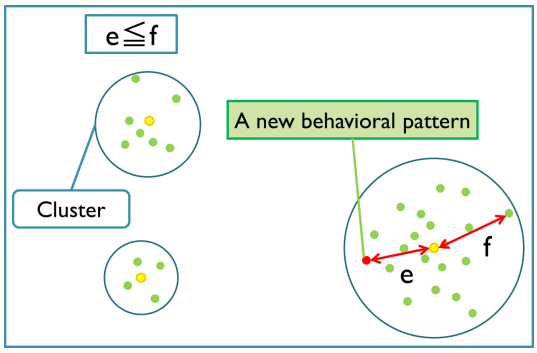


Figure 8. Normal situation

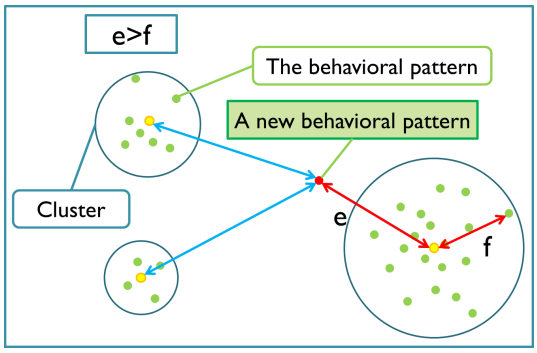


Figure 9. Irregular situation

in Section 3.1. Second, we calculate the Euclidean distances between the behavior pattern and each cluster center in 10-dimensional space. We extract the cluster with the smallest Euclidean distance, and denote the distance as e. We also denote the Euclidean distance between the cluster center and the most distant behavior pattern as f. If e is less than or equal to f, as shown in Figure 8, the observed behavior pattern is recognized as belonging to the cluster. If e is greater than f, as shown in Figure 9, the observed behavior pattern is considered to be an irregular situation.

4 Experiments and discussion

4.1 Experimental conditions

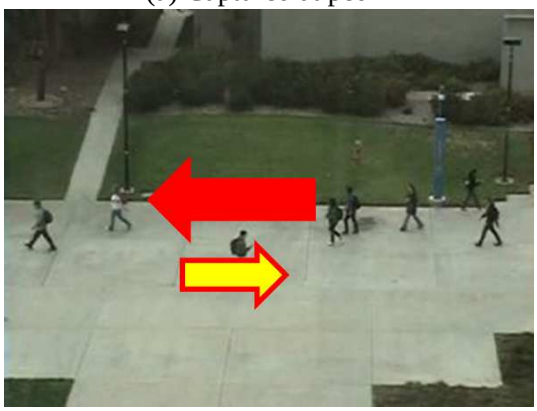
In order to verify the effectiveness of the proposed method, we performed an experiment using the ped 1 and ped 2 sample sequences in the UCSD dataset [6]. As shown in Figure 10, these are 320×240 [pixel] images captured from two different places. The presence of objects such as cars and bicycles is defined as an abnormal section in the UCSD dataset. There were 51 ab-

(a) Sample sequence ped 1

(b) Sample sequence ped 2

Figure 10. Example of sample images

(a) Captured at ped 1



(b) Captured at ped 2

Figure 11. Clusters containing behavior patterns captured from different locations

normal sections in the test dataset. First, we extracted behavior patterns from a training dataset that contained no abnormal sections. Following that, we learned and recognized the behavior patterns via the process of classification. Second, we extracted behavior patterns from a test dataset that did contain abnormal sections. Finally, we determined new behavior patterns to be "normal" or "irregular" by reference to the obtained clusters.

4.2 Experimental results

First, we extracted 82 behavior patterns from the training dataset, and classified them into 27 clusters. The clustering result is shown in Table 1. As can be seen from Table 1, we could generate six classes that were contained in the two different scenes. Second, we obtained 58 behavior patterns from the test dataset. Further, when a new flow was observed, we considered it to be an irregular situation. There were 19 behavior patterns classified as irregular situations. Of these, 14 patterns were contained in abnormal sections.

4.3 Considerations

4.3.1 Learning of behavior patterns

We classified images captured from different locations, from which we could obtain clusters containing behavior patterns. As shown in Figure 11, there were times when similar behavior patterns were classified into the same cluster. But as shown in Figure 12, sometimes when similar behavior patterns were not classified into same cluster. As shown in Figure 13, the frequency and norm were not extracted particularly well. This is thought to be because the sequence length of the behavior patterns was too short. Thus, we need to consider setting the optimal sequence length and normalization of the feature vectors.

4.3.2 Detection of irregular situations

As shown in Figure 14, when correctly detected, irregular situations contained objects such as a car or fewer people. When the detection was incorrect, as in Figure 15, there were

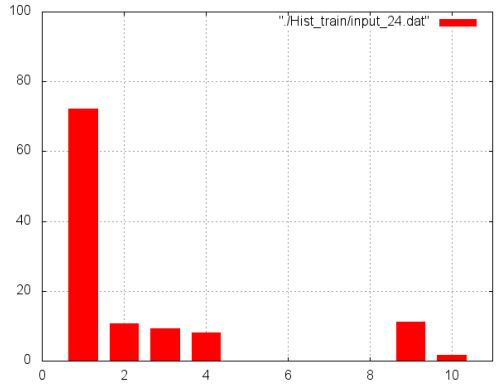
Table 1. Result of clustering

images	only ped 1	only ped 2	both ped 1 and ped 2
name of cluster	0,1,4,5,8,9,10,11, 12,13,14,15,16,17, 20,21,22,23,24,25	27	2,3,6,7,18,19
number of clusters	20	1	6

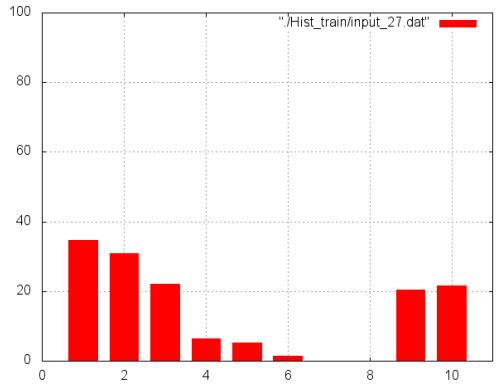
(a) Behavior patterns of No.24

(b) Behavior patterns of No.27

Figure 12. Images of similar behavior patterns are contained in the different clusters



(a) Behavior patterns of No.24



(b) Behavior patterns of No.27

Figure 13. The frequency and norm of similar behavior patterns are contained in the different clusters

generally many people. This is because the feature vectors were extracted according to the weight of all similar directions of movement. Therefore, we need to focus on features where the movement is large, even if the size of the object is small.

4.3.3 Intercomparison experiment for detection

We performed an experiment to provide performance measurements among our method and Mahadevan et al. method [4] using same UCSD dataset [6] which we used for experiments of detection in above sections. When their method demonstrated a about 90% true positive rate, the false positive rate is about 50%. In our method, we demonstrated a 27% true positive rate, the false positive rate is about 10%. If we use Mahadevan et al. methods, the behavior patterns must be relearned when the number of surveillance cameras is increased or their position is changed a little. But our method can detect abnormal behaviors using the same training data.

5 Conclusion

In this paper, in order to undertake a common learning process for different situations, we ex-

Figure 14. Example of correct detection

Figure 15. Example of false detection

tracted and classified feature vectors. When new behavior patterns were obtained, we extracted irregular situations with reference to predefined clusters. Further, in order to resolve the relearning problem, we used features without any location-based information. The clusters contained similar behavior patterns, even if they were from different moving images. This allowed us to confirm that the clusters were movement-based effects rather than artifacts of the location from which the video was taken. Our future work will consider features for which the movement is large, even if the size of the object is small, and the optimal sequence length.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 24700195.

References

[1] C. Stauffer and K. Tieu, "Automated Multi-Camera Planar Tracking Correspondence Modeling," IEEE Proceedings of Computer Vision and Pattern Recognition, pp. 259-266, 2003.

[2] L. Lee, R. Romano, and G. Stein, "Monitoring Activities from Multiple Video Streams: Establishing a Common Coordinate Frame," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), Vol. 22(8), pp. 758-767, 2000.

[3] S. Saxena, F. Bremond, M. Thonnat, and R. Ma, "Crowd Behavior Recognition for Video Surveillance," Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 08), 2008.

[4] Vijay Mahadevan, Weixin Li, Viral Bhalodia, Nuno Vasconcelos, "Anomaly Detection in Crowded Scenes," IEEE Proceedings of Computer Vision and Pattern Recognition, 2010.

[5] T. Gibo, S. Aoki, T. Miyamoto, M. Iwata, and A. Shiozaki, "Sequential Learning and Recognition of Comprehensive Behavioral Patterns Focused on Confluence," Proceedings of Japan–Cambodia Joint Symposium on Information Systems and Communication Technology (JCAICT2011), pp. 83-88, Jan. 2011.

[6] "UCSD Anomaly Detection Dataset," http://www.svcl.ucsd.edu/projects/anomaly/dataset.html

Consideration of page dwell time in B-to-B (Business-to-Business) industry web analytics

Akiyuki Sekiguchi, Kazuhiko Tsuda†

†Graduate School of Systems and information Engineering of Tsukuba

Abstract

Web analytics are basically used for improving usability and for marketing related uses.
Adding page dwell time to page view as a typical key factor in a B-to-B manufacturer's site analytics offers new insight. Although we have technical issues, we can see it is effective and deserves further study. In particular, page dwell time is very related to various types of conversion in a B-to-B site. Also, when we think about website structure improvement, looking at page dwell time can be one good metric.

Keywords: Web access analytics, page dwell time, web KPI

1 Introduction

Many B-to-B companies in the electrical industry have traditionally focused on large companies as customers. However, the large-customer market has become saturated (especially in Japan), and support for small and medium-sized companies, including the long tail market, must be considered for further business expansion. The strategy of assigning sales staff and performing much targeted marketing, which worked for large companies, cannot realistically be used for small and medium-sized companies due to support costs.
Therefore, it is critical that in addition to using a website as a medium for providing information to small and medium-sized companies, it should also be used for marketing activities such as collecting user information and getting a handle on the needs of the market. Furthermore, the volume of product information is increasing with the increasing functionality of products. And due to increasing numbers of customized products, the number of webpages and database records is also ever-increasing. Therefore, access analytics need to be used to optimize user paths through a website and so on. Moreover, there is a need to use analytics data for business (especially marketing and sales activities).
The research objectives of this study are the following. Typically page view and user visits are major **K**ey **P**erformance **I**ndicators (KPI) in web access analytics and page flow incorporating these are often used. However it is hard to see real analytics with an individual KPI like page view because pages with many page views might be caused by users confused by page navigation. There is very little study of combination of conversion rate and dwell time on pages and we tried to track the combination of dwell page and other indicators. This study is focused on B-to-B manufacturer business websites and the major characteristic is that the user goal for using a website is not always "to buy a product" but sometimes it is just to read articles or download files. All web analytics data comes from Renesas Electronics' Japanese website, from October 2012 to December 2012.

2 Overview of Purpose of Dwell Time on Pages

2.1 Preceding studies

There are many studies on web metrics like Ichikawa *et al*[1] and Xiao-Gang[2], and studies on web metrics related to the B-to-B market[3]. However, even if typical KPIs are discussed there is almost no study of the effectiveness of page dwell time on KPI.

2.2 Typical basic KPI

Web analytics are basically used for improving usability and for marketing related uses. Typically the following metrics are often used.

1) Page view
2) Unique users
3) Visits per user
4) Conversion rate

Page view is the most often used but it is hard to see true user flow with a single metric. One example is when we have many pages reaching a

final target page which is called the conversion page (conversion from "visitor" to "purpose achieved"); total page view count is high, but most pages could be unnecessary for users. In this case we need to look at page dwell time as well as the page view metric. Presumably a combination of page view and page dwell time can be an important metric set for measuring user visit behavior. On top of that site dwell time and page dwell time is often discussed in web analytic academic studies, but there are very few studies on the combinations.

2.3 Overview of page dwell time

Firstly we need to define what page dwell time is. Page dwell time is the length of time that users stay on each page. Note that it is different from site dwell time.
The similar metric "site dwell time" is the length of time that elapses from the first moment a web user enters a particular website until the time that user leaves that website.

2.4 Expected benefit from page dwell time analytics

The following benefits are expected when we implement tracking of page dwell time.

1) Combination of page view and page dwell time can allow more detailed analytics of page role and the gap between design and results.
2) It helps categorize the page type, like index page or article page, which users actually read through.
3) Possible to evaluate page value related to user conversion with dwell time, and score pages for each conversion.

2.5 Technical issues on page dwell time tracking

The following items are open issues and we need to consider them when we analyze data.

1) The last page when users exit cannot be tracked at all due to web beacon and http protocol characteristic.
2) When new tab in browser is used for clicking the captured data can be wrong because user may keep it alone in this tab and browse pages in different tabs.
3) In general some users keep pages open pages for a while even when not viewing them. So we need to remove some abnormal length data.

We need to take care of these items when we check data.

3 Consideration of the relationship between dwell time and page view

In a B-to-B manufacturer's site, assumed page categories are the following.

1) Index type page: Kind of a table of contents of child pages.
2) Product catalog type page: Contains lots of information in a standardized format and is often used.
3) Article type page: Contains an article. Page length varies greatly.
4) Transit type page: One of the steps in a procedure, like user registrations.
5) Conversion page. This is a final destination page and can differ depending on the purpose of the user. Sometimes it can be a purchase page and other times it can be a download page.

3.1 Product catalog page analysis.

In this product family section we have over 200 pages and each page contains information, which can be a volume of one A4 size page when printed.
Firstly we need to review the overview of the dwell time data. Figure 1 shows the relationship between page view and page dwell time. We see most users stay on a page for about one minute.
Page dwell time of over two minutes is not often seen, and even when we do it is with very few page views.

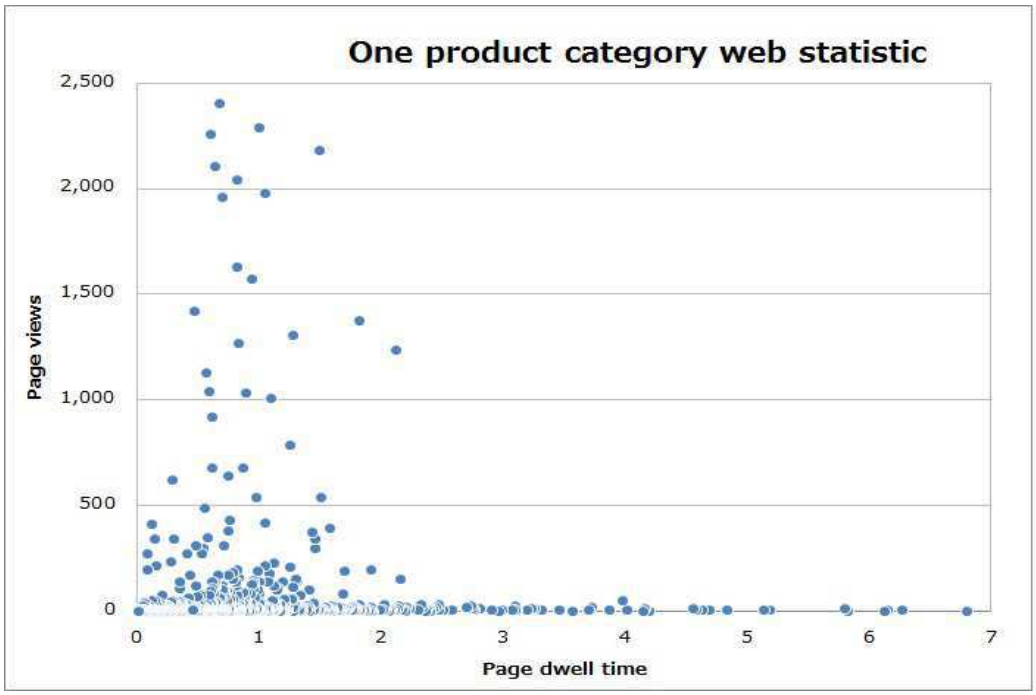


Figure 1. Typical relationship between page view and dwell time

3.2 Trends for article type page and index type

The following are the results of metrics on pages which are designed for reading, like FAQ (Frequent Asked Question). There are two types of FAQ pages. One is index type which consists of lists of FAQs with a short description of each FAQ and the other type is article, which is an actual FAQ page. Typically the information volume of an actual FAQ page is two or three pages of A4 paper.
In fact an actual FAQ page has longer dwell time than an FAQ index page. The dwell time of the index type is 0.9 minutes average and the article type is 1.5 minutes average. Please refer to Figure 2. Page dwell time is dependent on page type.

4 Relationship between dwell time and conversion participation

4.1 Dwell time of contribution pages to conversions

'Conversion' means a user's final goal in web activity. In the typical B-to-B market area the conversion target of a web site isn't necessarily a purchase. Download of an object such as program source or a document, and a sales inquiry can be conversions sometimes. Especially, user registration for the website is the most common conversion. Considering the relationship between conversion and page dwell time we checked the dwell time of pages passed through by users to the destination page, which brings us to conversion. We call these pages "conversion related pages". We had originally expected a conversion related page to have longer dwell time because these pages are related to user decision. However, looking at the statistics we can say that the dwell time of conversion related pages isn't longer than the average for all pages. It seems that users generally go quickly across pages into final conversion. In particular, even user registration pages, which are mostly commonly used by B-to-B customers, don't have longer dwell times. Please refer to figure3. The page that contributed the most (product/mpumcu/index.html) has 0.42 min. as the average dwell time. This is much shorter time than the general average. This shows that users pass through pages in a short time and finally reach the goal page.

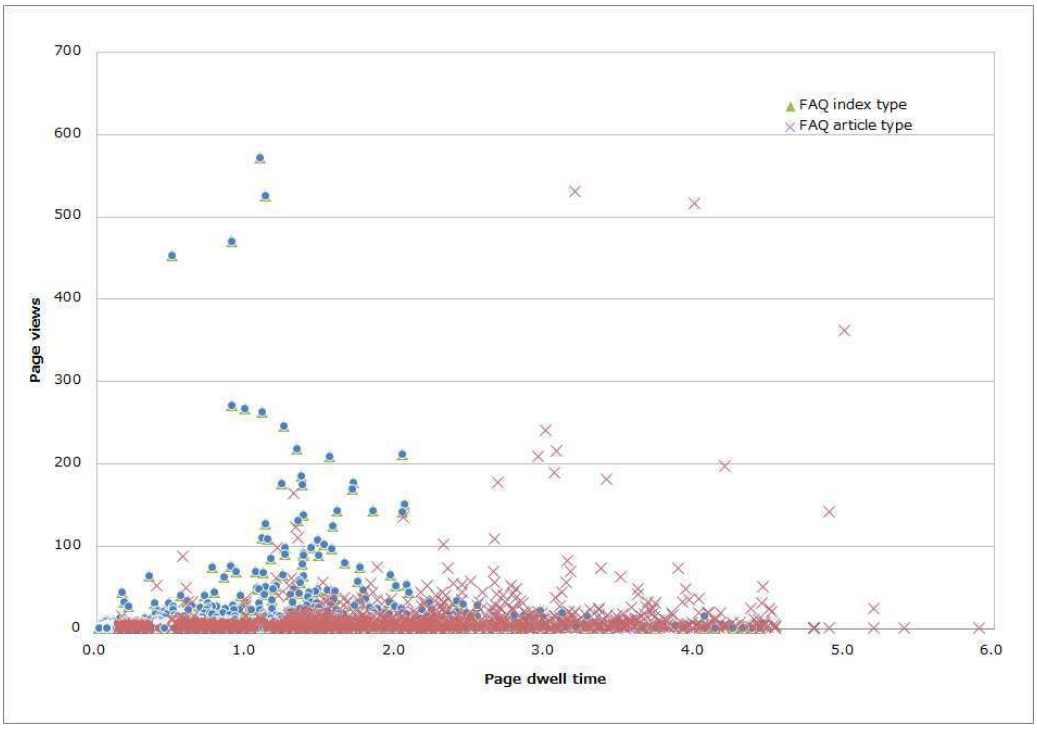


Figure 2. Dwell time on index type page and article type page

Page name	Dwell time (min.)	Page view	Product lineup	Parametric search	Registration	Buy
products/mpumcu/rl78/index	0.46	30,250	1,378	2,738	255	601
products/mpumcu/index	0.42	46,478	4,450	3,767	575	599
products/mpumcu/rx/index	0.37	32,978	1,182	1,855	220	450
products/mpumcu/r8c/index	0.41	21,193	1,080	2,227	205	387
products/mpumcu/superh/index	0.53	30,912	1,042	2,475	349	347
products/mpumcu/rl78/rl78g1x/index	0.31	8,399	360	992	29	285
products/mpumcu/rl78/rl78g1x/rl78g13/index	1.04	8,111	459	1,177	31	262
products/mpumcu/rx/rx600/index	0.50	14,402	330	932	143	234
products/mpumcu/rl78/rl78g1x/rl78g14/index	0.96	4,012	201	734	10	232
products/mpumcu/rx/rx600/rx621_62n/index	0.91	9,200	181	668	54	212
products/mpumcu/rl78/rl78g1x/rl78g12/index	0.80	3,545	210	805	31	177
products/mpumcu/78k/index	0.51	10,848	1,281	1,285	205	175
products/mpumcu/r8c/r8c3x/index	0.49	7,601	263	887	34	160
products/mpumcu/h8/index	0.47	11,186	750	685	381	135
products/mpumcu/v850/index	0.50	12,611	797	743	118	130
products/mpumcu/rx/rx200/index	0.51	8,494	348	568	52	130
products/mpumcu/superh/sh7080/sh7080/index	0.83	1,720	74	284	21	112
products/mpumcu/superh/sh7216/sh7216/index	0.85	2,231	68	377	1	100
products/mpumcu/r8c/r8c2x/index	0.33	3,277	114	436	36	97
products/mpumcu/rx/rx200/rx210/index	0.96	6,787	306	621	32	95

Figure 3. Dwell time on conversion related pages (top 20 by registration conversion)

There are 2627 unique pages related to conversion in this data set.. Please refer to Table 1 and Figure 3.

Table 1. Overall statistics of dwell time on conversion related page

Unit: minutes

No.	Conversion related page	Average time	Maximum time	Minimum time
1	Product lineup list page	1.01	2.29	0.35
2	Parametric search results	0.99	2.38	0.09
3	User registration	1.01	2.51	0.09
4	Buy	0.84	2.24	0.08
5	All pages related to conversion	0.95	3.02	0.03

Also "buy through distributors" is one of the B-to-B market conversions. We expected the pages related to "buy through distributors" conversions to have longer dwell time than the total average (around 1 min.) because this conversion could take some decision-making. However, this average time for this conversion is 0.84 min., which is shorter than the total average. This is one of the characteristics seen with B-to-B sites– we assume users had already decided to buy product due to a previous or earlier visit to the website so when they purchase they go straight to the purchase without examining product specifications.

We checked the pages that were used the most for conversions. In the above Table 1, no. 2 and no. 4 were the most important conversions. For user registration conversion, software download pages are highly ranked. They make up 34 pages among the top 100 pages and the next most numerous are sample code/document download pages. This is understandable for a B-to-B manufacturer's website.

Also for no. 4 "Buy" conversion, the most common conversion related pages are product spec pages. They make up 20 pages among the top 100 pages.

As a result we can assume users examine the product specifications even if only for a short time.

4.2 Dwell time of only previous pages to conversions

In section 4.1, we did an analysis for all conversion related pages. Next, we focus on only the previous pages to conversions. In 4.1, generally we can say that dwell time of all conversion related pages is not so long compared with the general average, but we need to check the time for previous pages only.

As we stated earlier in this report one important conversion is user registrations. When we look at the dwell time data for previous pages that were used the most, it is 1.79 min. average, which is longer than the general average. We assume many users select the proper downloaded files carefully before performing user registration. Also one note is that the second most viewed page is the explanation page for user registration and the disclaimer. Its dwell time is 0.38 min. and relatively low. We see that most users pass through this page, but they don't take a long time with it. This is another one of our key findings.

The average dwell time of previous pages of conversions is 1.13 min. Our original expectation was longer than normal dwell time on the average. However, it is actually almost the same as other pages, even if the page comes very soon before a conversion.

Page name	User registration	Vsitors	page view	Dwell time (min.)
support/downloads/download_results/C100(	51	212	517	1.79
support/how_to_use_myrenesas/index	50	34	92	0.38
index	39	7,143	15,192	0.48
support/downloads/download_results/C100(	32	179	384	1.04
request	30	1,331	2,751	0.88
myrenesas/index	26	140	244	0.50
req/part_no_search	20	956	3,838	1.56
support/downloads/download_results/C100(	19	63	112	1.65
support/downloads/download_results/C2017	19	91	206	1.82
products/soc/usb_assp/document/registrati	18	4	16	0.53
products/tools/ide/cubesuite_plus/downloac	17	221	529	0.55
privacy/index	16	12	15	0.76
support/downloads/download_results/C200(	15	159	268	1.07
support/downloads/download_results/C200(	15	80	156	1.23
products/mpumcu/rx/rx600/rx62t/app_notes	13	51	119	1.42
support/downloads/download_results/C2017	12	17	35	1.32
products/tools/ide/hew/downloads	10	330	793	0.48
products/mpumcu/rx/support	10	84	111	0.71
products/tools/evaluation_software/downloa	9	257	643	0.39
products/tools/evaluation_software/index	9	390	740	0.63

Figure 4. Dwell time of previous page of user registration conversions

Page name	Vsitors	page view	Dwell time (min.)	Buy
products/tools/emulation_debugging/onchip	390	849	1.15	72
buy/disti/index	65	98	0.95	53
products/tools/emulation_debugging/onchip	239	536	1.04	35
buy/index	298	430	0.98	18
req/search	4,017	12,934	0.82	11
products/mpumcu/superh/sh7080/sh7080/ir	109	221	0.86	11
products/mpumcu/rx/rx600/rx621_62n/inde	603	1,164	0.88	11
products/tools/emulation_debugging/onchip	95	225	0.83	10
jump	76	119	0.67	9
req/part_no_search	956	3,838	1.56	9
products/mpumcu/rx/rx600/rx610/index	154	271	0.79	9
req/ecommerce	43	62	1.32	8
products/mpumcu/superh/sh7216/sh7216/ir	158	305	0.67	8
req/ecommerceTree_search	41	63	1.44	6
products/mpumcu/rl78/rl78g1x/rl78g13/inde	551	1,043	0.87	6
products/mpumcu/superh/sh7137/sh7137/ir	24	45	1.31	5
products/mpumcu/rl78/rl78g1x/rl78g12/inde	232	420	0.63	5
req/keyword_search	1,041	2,811	0.96	4
req/parametric_search	296	540	0.92	4
products/mpumcu/superh/shether/sh7619/ir	59	99	0.85	4

Figure 5 Dwell time of previous page of purchase conversions

Please refer to Figure 5. Even if the purchase is important and a big decision for users the average dwell time is 0.97 min., which is not long. This is the same trend as conversion related page time which was stated in section 4.1.

5 The relationship between text volume and dwell time

One potential factor related to dwell time is page content volume. Pages have different amounts of information and dwell time of users can depend on page size.
Page volume can be measured by text size in HTML (Hyper Text Markup Language). Refer to Figure 6. Generally speaking text size is not much related to dwell time. This could be a feature of a B-to-B type web site. Dwell time can probably be more related to page type or conversions. We will study the relationship of page type with dwell time in more detail in further studies.

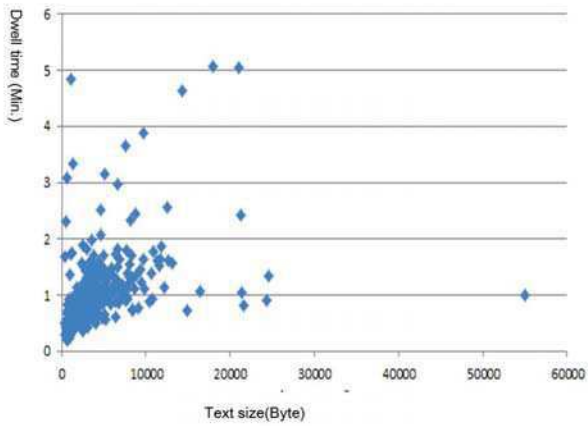


Figure 6. Relationship with dwell time in detail in further studies.

6 Conclusion

We examined the web access analytics adding page dwell time on page view as a typical key factor for a B-to-B manufacturer's site. Although we have technical issues, we can see it is effective and deserves further study. In particular, page dwell time is very related to various types of conversion in a B-to-B site. Also, when we think about website structure improvement, looking at page dwell time can be one good metric. We can test usability or expected page flow with the results of dwell time as well as page views.
As one possibility page length could be related to dwell time, but in our first general study we didn't see any relationship between page length and dwell time compared with other factors like page type or conversion contribution. We need to study this topic further. Also we will try examining the effectiveness of page reading completion rate with dwell time because we found many users quickly pass through pages until conversions.

7 For More Information

For more information or queries complying with the submission process, e-mail s1330180@u.tsukuba.ac.jp sufficiently before the submission deadline so that alternate arrangements can be made. For help with difficulties in electronic submission or technical problems, you can contact the above e-mail.

Acknowledgment

Thanks to Renesas Electronic Corporation. This format is a modification of KICSS2006 and IJCKS2007 style file.

References

[1] Y. Ichikawa, M Nakamura, Y. Kishimoto, T. Kobayashi, "A Proposal of Extracting Innovative Users with Web Access Log of an E-Commerce site" IPSJ SIG Notes 2012-GN-83(2), 1-7, March 2012.
[2] W. Xiao-Gang, "Web mining based on user access patterns for web personalization": Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on,pp.194-197, 2009.
[3] A practical evaluation of Web analytics.. *Emerald Article*, Internet Research, Vol. 14 Iss: 4 pp. 284 – 293,2004.
[4]Sampath, P.Dept. of Comput. Sci. & Eng., Bannari Amman Inst. of Technol., Sathyamangalam, India, "An efficient weighted rule mining for web logs using systolic tree": Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on, pp.432-436, March 2012.
[5] S.Otsuka, M.Toyoda, M.Kitsuregawa, "A Study for Analysis of Web Access Logs with Web Communitie": Information Processing Society of Japan (IPSJ), Database44, pp.32-44, 2003.

How Do Customers Buy Them at a Supermarket? Behavior Analysis from Real Observation and Agent Simulation

Masaki Kitazawa† **Fumiaki Sato†** **Takashi Yamada†**

Masakazu Takahashi‡ **Takao Terano†**

†Tokyo Institute of Technology, 4259-J2-52, Nagatsuda-Cho, Midori-Ku, Yokohama-Shi, Kanagawa 226-8502, Japan
{m.kitazawa, satof, tyamada}@trn.dis.titech.ac.jp, terano@dis.titech.ac.jp
‡Yamaguchi University, 2-16-1, Tokiwadai, Ube-Shi, Yamaguchi 755-8611, Japan
masakazu@yamaguchi-u.ac.jp

Abstract

This paper presents a customer behavioral model for grounding the number of purchase items in Agent-Based In-Store Simulator (ABISS). ABISS is a decision support system for a retail management. Using ABISS, we are able to virtually investigate the shopping paths of the customers and analyze the effect of sales promotion. To grounding the number of purchase items, first we conduct a field study using Radio Frequency Identification (RFID) and analyze the collected RFID data and Point-of-Sales (POS) data. Then we develop a decision model, which determined customers' "Shopping List", "Possession Money Limit", and "Staying time". The experimental result has revealed that we ground the number of purchasing items of obtained by the POS data, if we concentrate on the tuning of ABISS parameters on "Shopping List" and "Possession Money Limit".

Keywords: Agent-Based Simulation, Agent-Based Modeling, Customer Behaviors, Radio Frequency Identification

1 Introduction

Service management in supermarkets requires the analyses on the effects of sales promotion and the control of customer shopping paths in the store. So far, such analytic tasks cost so much in a real store, and it is difficult to measure the effects of sales promotion. To let them easily carry out such analyses, we have developed a decision support system with Agent-Based In-Store Simulator (ABISS), for retail management. Using ABISS, we are able to virtually investigate the shopping paths of the customers and analyze the effect of sales promotion (Figure 1)[1].

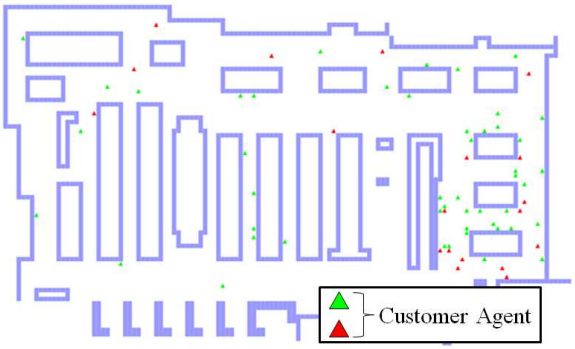


Figure 1. Snapshot of ABISS

To enhance the performance of ABISS, however, we must observe real customer behaviors and determine their behavioral parameters. Ordinarily, in a such analysis we must trace real shopping paths by human observers, which required very high costs and long time for the analysis. Instead, in this paper, we describe a new indirect measurement method: First, we utilize a general purpose RFID tag system equipped with both in market carts and shopping places; Second, purchased items are collected to the point of sales information gathered at the cashing register places; and Third, the gathered shopping paths with purchase information are compared with intensive simulation results by ABISS.

This paper presents the customers' behavioral model for grounding of the number of purchasing items in a supermarket and experimental result of the simulation. First, we conduct a field study using RFID and analyze the collected RFID data and Point-of-Sales (POS) data to build the model. Second, we build a decision model in which determined "Shopping List", "Possession Money Limit", and "Staying time". The result has re-

vealed that we ground the number of purchasing items of obtained by the POS data, if we consider only “Shopping List” and “Possession Money Limit” in ABISS.

The rest of the paper is organized as follows: Section 2 discusses the background of the research and related work; Section 3 explains the field study in the target store; Section 4 describes the setups of ABISS; Section 5 presents experimental results of ABISS; and Section 6 gives concluding remarks and future work.

2 Background and Related Work

A retail store performs many sales promotions in order to make sales increase. However, such sales promotion cost so much in a real store. Then, in order to perform sales promotion effectively, there have been so many researches about customer behavior in marketing science. Conventionally, many data mining is used [2], [3]. Moreover, flow line analysis of the customer using an RFID tag or accelerometer has been also used [4], [5] and [6]. In recent years, the method of predicting the effect of sales promotion by Agent-Based Simulation (ABS) has attracted attention. In the study of supermarket, we usually model the “Agent” buying behavior and walking behavior of a store. In order to predict the effect of sales promotion, it becomes important to have grounded with the number of purchases with the actual number of purchases of imagination. In order to ground the number of purchases, it is required to take the following into consideration: 1) the probability which purchases each goods, 2) the conditions which customer stops shopping. The probability which purchases each goods is calculated from POS data in many cases. On the other hand, the conditions which a customer stops shopping is seldom considered.

The conditions which a customer stops shopping in ABS is shown in Table 1. In [7], [8], [9], and [1], if Customer Agent purchases all the goods on the shopping list possessed beforehand, the model which stops shopping will be adopted. Moreover, in [9] and [10], the missing item from the side of the prices of goods, length of stay, and a store is also taken into consideration as conditions which stop shopping.

ABISS is one of the simulators for predicting the customer's of the supermarket in which we have been developing several years. ABISS carries out the virtual experiment of customers’ purchasing and shopping paths in the supermarket. In ABISS, the model in consideration of the budget of a shopping list and a customer is adopted. In ABISS, the agent's parameter has been set based on the real POS data. Also, so far, we have experimented in change of the flow line length and the layout change. However, the customer's buying behavior model was not considered by the viewpoint of grounding of the number of purchasing items. In order to clarify the characteristics of the behavior model, this research investigates the conditions on which the customer agent of ABISS stop shopping in the sense of purchasing items. For the purpose, first, the real data on a customer's length of stay are acquired in the inside-of-a-shop experiment using an RFID tag system. Next, the limit of money possession with the staying time, the shopping list calculated from the POS data, and customer decision rules are set as ABISS parameters.

Table 1. The Conditions Which a Customer Stops Shopping in ABS

	Customer Reason				Store Reason
	Shopping List	Possession Money Limit	Staying Time	Purchased Weight	Stockout
Hayashi [7]	○	×	×	×	×
Miwa [8]	○	×	×	×	×
Macal [9]	○	×	○	×	○
Siebers [10]	×	○	○	×	○
Terano [1]	○	○	×	×	×

3 Field Study in the Target Store

3.1 Setup

The condition of target store and setups of field study is summarized in Table 2. The store layout and the placement of RFID antenna which ABISS targets are shown in Figure 2. An order of vegetable, fish, and meat category from the entrance is widely employed in Japan. In this store, we set 10 RFID antennas on display fixtures, and 6 RFID antennas on each registers. By the antenna arrangements, we are able to measure the whole cart moves inside of the store. The same RFID tags are also set in the 53 shopping carts. The antenna setup examples are shown in Figure 2. This field experiment was conducted from 9:00 to 21:00 on March 3-5, 2011.

Table 2. Condition of Target Store and Setups of Field Study

Location	Shimane Prefecture, Japan
Opening Time	9:00 – 21:00
Average of Customers Per Day	2000 customers
Experimental Period	March 3 – March 5, 2011
RFID Antennas	10 on display fixture 6 on each register
RFID Tags	53 on shopping cart (Half of shopping cart)

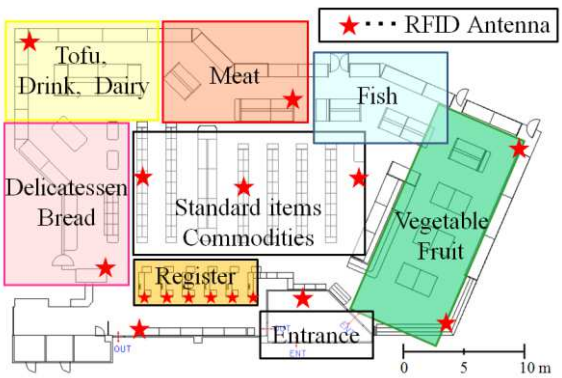


Figure 2. Store Layout and Position of RFID Antenna

3.2 Result

We analyze the RFID and POS data on March 4, 2011. Since sales promotion of the day was less than those in the other days, we are able to observe basics of the customer behaviors. The total number of RFID data on March 4 is 102,661 in the sense of the information communications. The POS data on March 4 is about 1,849 customers and 1,882 numbers of purchasing items. The analysis is conducted in the following procedures. First, we identify the RFID data for every customer. A customer's store entrance time is measured by the time of RFID of the Entrance and the antenna of a cart communicating at the first time. Also, a customer's store exit time is determined as the time of communicating at the antenna of the last of a register place. Second, we

(a) RFID Antenna on Display Fixture (attached the pink cover)

(b) RFID Antenna on a Shopping Cart (banded with white string)

Figure 3. Setup of RFID Antenna

give the relation of the POS data and the RFID data of each customer. That is, such data of the correspondence as register passage time and a customer's exit time are extracted. Accordingly, we get 100 data in which the customer's purchase items and shopping paths are related explicitly, although these data are independently collected. Since the other data have so many noises, we are not able to relate them correctly. The amount of key statistics of a customer's amount of money for purchases and their staying time are shown in Table 3. Moreover, scatter diagram of the histogram of a customer's amount of money for the purchase, the histogram of stay time, the amount of money for purchase, and stay time are shown in Figure 4.

The correlation values of a customer's amount of money for purchase and stay time are -0.007. It is a very small value, and the analysis which connected RFID Data and POS Data showed that they were independent.

4 Setups of ABISS

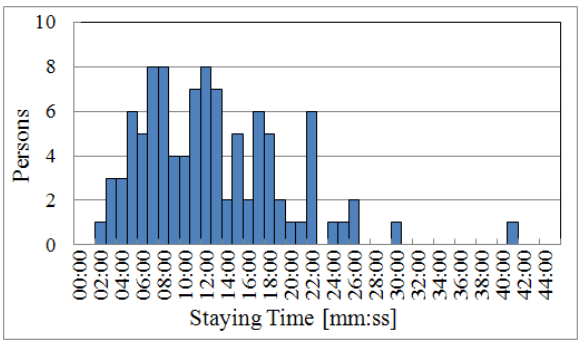
Based on the result in Section 3, we prepared three customers' buying behavioral models which are shown below and by which the customer changed the conditions which stop shopping. The flow figure of each model is shown in Figure 5.

1) Shopping List & Money Model
In this model, customer agents finish shopping when a shopping list become empty or the amount of purchase money exceeds a budget.
2) Shopping List & Time Model
In this model, customer agents finish shopping when a shopping list become empty or the staying time exceeds a schedule.
3) Shopping List, Money & Time model
In this model, customer agents finish shopping when a shopping list become empty, the amount of purchase money exceeds a budget or the staying time exceeds a schedule.

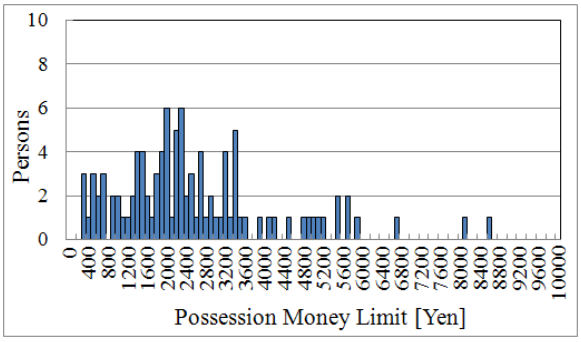
We calibrate the parameter from real data. The number of planned items and distribution of purchase money are calculated from the POS data on March 4, 2011. Also, the distribution of staying time is calculated from RFID data on March 4, 2011. The kind of parameters other than the above is the same as in [1].

Table 3. Basic Statistics of Money and Time

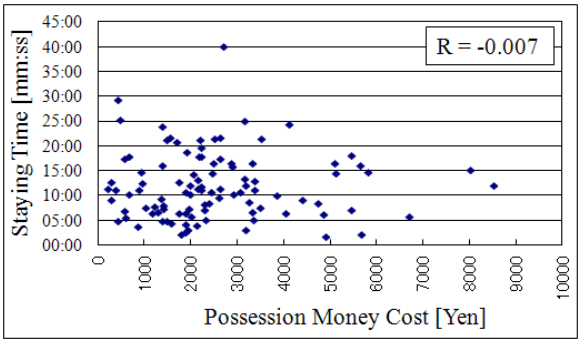
	Possession Money	Staying Time
Average	2524	0:12:01
Std Dev.	1643	0:06:45
Median	2208	0:11:07
Mode	1394	0:10:12
Min	204	0:01:51
Max	8529	0:40:13
N	100	100
R	-0.007	



(a) Histogram of Possession Money Limit

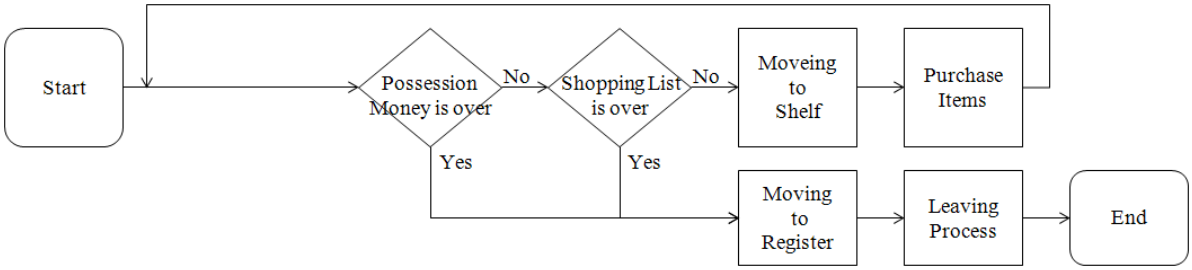


(b) Histogram of Staying Time

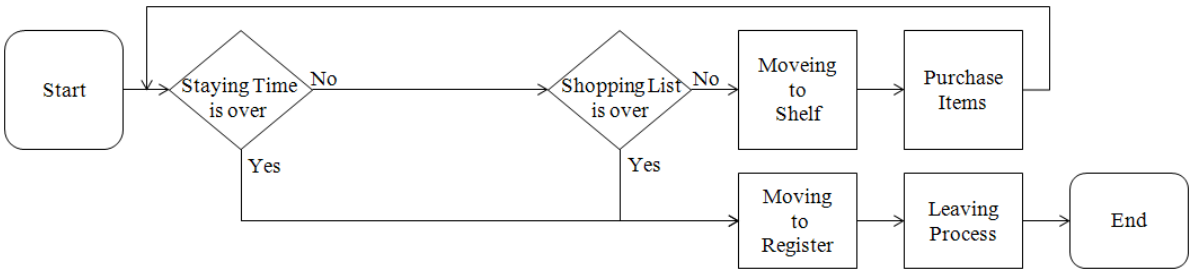


(c) Scatter Diagram of Staying Time and Possession Money Limit

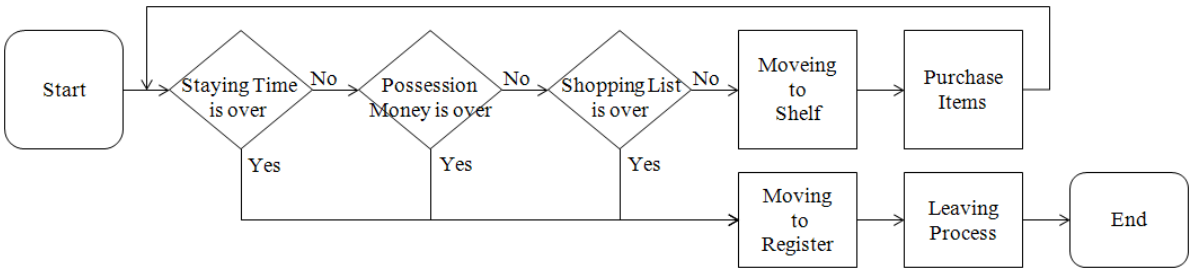
Figure 4. Results of Analyzing RFID data and POS data



(a) Shopping List & Money Model



(b) Shopping List & Time Model



(c) Shopping List, Money & Time Model

Figure 5. Flow Diagram of Customer Behavioral Model

Table 4. Experimental Settings on Customer Behavioral Model

Customer Behavioral Model	・Shopping List & Money Model ・Shopping List & Time Model ・Shopping List, Money & Time Model
Simulation Time	9:00 - 21:00 (43,200 simulation steps)
No. of Simulation Runs	3 times

5 Experimental Results of ABISS

Based on the model Section 4, we have conducted simulation experiments. We have conducted five 43200-step (representing the store's opening time) simulations. The experiment settings are summarized in Table 4. We have analyzed how the changes in model affected the customers' purchased number.

Table 5 shows number of real and simulated purchased item number of customers. We adjusted each experimental result so that the number of samples and real data may become the same. An error rate shows a difference with the number of purchasing items in the POS data on March 4, 2011. We have found that the Shopping List & Money Model gives the smallest error rate. It suggests that we can ground the number of purchasing items in ABISS of POS data even if we do not consider "Staying Time". It seems that the noise, for example conversation time with friends in a supermarket, is mixed with the data of "Staying Time".

Table 5. Experiment Number of Purchased Items Tuned to Real Transaction

	Samples	Purchased Number		Error Rate [%]
		Average	Std Dev	
POS	1849	18882		
Shopping List & Costs	1849	18513	243	-1.95
Shopping List & Time	1849	19503	4575	3.29
Shopping List, Costs & Time	1849	15845	257	-16.08

6 Conclusion

This paper has presented a customer behavioral model for grounding the number of purchase items in Agent-Based In-Store Simulator (ABISS). We have described the setups and the analyzing results of field study, and the models which changed a stopping condition of shopping and experimental results of ABISS. First, we have carried out field study with RFID tags and analyzed staying time. Second, we have made a model in ABISS which changed a limit of shopping. We have decided the limit from "Shopping List", "Staying Time", and "Possession Money Limit". The experimental results have suggested that we can ground the number of purchasing items in ABISS to POS data if we consider only “Shopping List” and “Possession Money Limit”.

Our future work includes 1) analysis of customers’ branching probability in a real store to ground walking flow, 2) development of decision support functions for store managers, and 3) analysis regarding customer information and product categories.

References

[1] Terano, T., Kishimoto, A., Takahashi, T., Yamada, T., Takahashi, M.. Agent-Based In-Store Simulator for Analyzing Customer Behaviors in a Super-Market. In: Velasquez, J.D., Rios, S.A., Howlett, R.J., Jain, L.C. (Eds.) KES 2009, Part II, LNCS (LNAI), vol. 5712, pp. 244-251, Springer, Heidelberg, 2009

[2] Yada, K., Motoda, H., Washio, T., Miyawaki, A.. Consumer Behavior Analysis by Graph Mining Technique, New Mathematics and Natural Computation 2, 59-68, 2006

[3] Yada, K.: String analysis technique for shopping path in a supermarket. Journal of Intelligent Information Systems, Vol. 36, No. 3, 385-402, Springer, Heidelberg, 2009

[4] Thiesse, F., Al-Kassab, J., Fleisch, E.. Understanding the value of integrated RFID systems: a case study from apparel retail, European Journal of Information Systems, vol. 18, 592-614, 2009

[5] Larson, J.S., Bradlow, E.T., Fader, P.S.. An Exploratory Look at Supermarket Shopping Paths. International Journal of Research in Marketing, vol. 22, no. 4, 395-414, 2005

[6] Kitazawa, M., Takahashi, M., Yamada, T., Yoshikawa, A., Terano, T.. How Do Customers Move in a Supermarket? -Analysis by Real Observation and Agent Simulation. The 3rd Japan-China Joint Symposium on Information Systems (JCIS2010), Beijing, China, 2010

[7] Hayashi, H., Cho, K.. Evaluating In-Store Digital Signage by Agent-Based Simulation. 3rd World Congress on Social Simulation (WCSS2010), Kassel, Germany, 2010

[8] Kanna Miwa,Soemon Takakuwa. SIMULATION MODELING AND ANALYSIS FOR IN-STORE MERCHANDIZING OF RETAIL STORES WITH ENHANCED INFORMATION TECHNOLOGY. Simulation Conference,: 1702-1710, 2008

[9] Charles M. Macal, Michael J. AGENT-BASED MODELING AND SIMULATION: DESKTOP ABMS. Simulation Conference, 39:95-106, 2007

[10] Peer-Olaf Siebers, Uwe Aickelin. AN AGENT-BASED SIMULATION OF IN-STORE CUSTOMER EXPERIENCES. Operational Research Society 4th Simulation Workshop (SW08). 2008.

Ontology-Enhanced Life Cycle Assessment: A Case Study of Application in Oil Refinery

Akkharawoot Takhom[†] **Boontawee Suntisrivaraporn**[†] **Thepchai Supnithi**[‡]
[†]School of Information, Computer and Communication Technology,
Sirindhorn International Institute of Technology, Thammasat University, Thailand
[‡]Language and Semantic Technology Laboratory, Intelligent Informatics Research Unit,
Intelligent Informatics Research Unit,National Electronics and Computer Technology Center, Thailand
akkharawoot.t@gmail.com

Abstract

This paper presents an improvement of a previous work on an ontology for life cycle assessment or O-LCA by taking into consideration the life cycle inventory (LCI) and life cycle impact assessment (LCIA). The knowledge about environment, manufacturing processes, participating substances, as well as that about assessment methods, are represented in the web ontology language. To demonstrate the benefits of employing a logical formalism and inference in life cycle assessment, we implement a resource recommender system built on top of our developed ontology, and carry out a usability study of the system. With respect to a case study of oil refinery, the evaluation suggests that though there are still a number of requirements from the field users, the system was proven useful in term of user satisfaction.

1 Introduction

To measure an environmental impact, a kind of environmental impact assessment (EIA) tools called *Life Cycle Assessment (LCA)* is used to identify, quantify energy and materials which used and released to the environment, and evaluate and implement opportunities to influence environmental improvements. To utilize the knowledge, the *Web Ontology Language (OWL)* [1], endorsed by the World Wide Web Consortium (W3C), is applied to improve the knowledge domain, called *O-LCA ontology* [2]. The *Description Logic (DL)* [3] manifests its capability to formalize the LCA ontology, allowing us to represent and reason with conceptual knowledge [4]. Automatic reasoning of DL is tractable as it is a tractable subset of first-order predicate logic, resulting in expressivity and robust scalability.

In previous work of LCA knowledge and semantic web technology, the standard guideline has been used to interpret and adapt with the specific objectives. Some of related work focuses on utilization of different standard guidelines. For example, CASCADE [5] project aims to adapt the data document format with other standard guidelines. LCAO [6] is another project that interprets the LCA framework, and aims at the organization and retrieval of information, and at the contribution for a consensual vision. In the case study of U.S. energy impact data management [7]. LCI can also be semantically represented as manipulatable databases using relational algebra. In this approach an ontology is used to model associate elementary processes in LCI according to their interdependency relations.

For enhancing the formalized LCA knowledge, this work carries on improving the $\mathcal{O}_{\mathrm{LCA}}$ [2] by expanding LCA phase to *Impact Assessment (LCIA).* The renovated knowledge is implemented as a recommender system by $\mathcal{O}_{\mathrm{LCA}}$ *Framework* and applied to a case study of oil refinery. By working proof, an inferential ability of ontology based on DLs system is utilized in the recommender system that follows the approach of *Cleaner Technology* [8] for environmental impact reductions.

The remainder of this paper is as follow. We will start with background knowledge in both of knowledge representation and Life Cycle Assessment as the domain knowledge. Next section, an implementation with $\mathcal{O}_{\mathrm{LCA}}$ Framework will be described: redesigning enhanced the ontology to support the framework, and remodeling ontology from exists to a new $\mathcal{O}_{\mathrm{LCA}}$. A case study of oil refinery is chosen and adjusted to be suitable usage in the fourth section. With a purpose of the practical usage, important issues will be discussed in the fifth section. Finally, we will

conclude and explain future work in last section.

2 Background

2.1 Semantic Web Ontology Language

Ontology is a computational artifact for encoding knowledge inferred from the domain of interest. The ontology formalizes tacit knowledge and domain expertise by explicitly representing them with a set of concepts, i.e. conceptualization, and relation elicitation. To code ontology, the use of OWL [1] is endorsed by the World Wide Web Consortium (W3C), because the language is based on semantic web and RDF/XML serialization.

To express the knowledge, DL [3] helps formalize the LCA ontology and allows us to represent and reason with conceptual knowledge. OWL manifests adequate expressivity and guaranteed tractable reasoning runtime. It features several kinds of axioms that are equivalent to *concept inclusion (CI)* and *role inclusion (RI)* of DL as shown in Tab. 1. Concept definition (CD) $A \equiv D$ is equivalent to two concept inclusions (CI) $A \sqsubseteq D$ and $D \sqsubseteq A$. There are two components for storing the knowledge: *TBox (terminological box)* and *ABox (assertional box)*. TBox contains concepts, concept hierarchies (encapsulated in terms of role inclusion), relations, and relation hierarchies. ABox represents relationship among individuals and concepts (via concept assertion) and that among individuals (via role assertion), as shown in Tab. 1. The last component, reasoning, is utilized to infer logical consequences from a set of asserted axioms.

The *DL-based Web Ontology Language (DL systems)* is popular in the semantic web context. Different DL systems offer different reasoning services, as known as inference problems, to explicate certain implicit knowledge logically captured in an ontology. A class of reasoning services are commonly considered mandatory and supported by most DL systems. For the example class of reasoning services, the class hierarchy that is automated by a reasoner, called the inferred hierarchy, which is used to checking the inconsistent classes.

2.2 Life Cycle Assessment

Life Cycle Assessment (LCA) [9] is a technique for measuring impact associated with a product, a process or an activity. This technique is to identify and quantify energy and material uses and releases to the environment, and to evaluate and implement opportunities to influence environmental improvements.

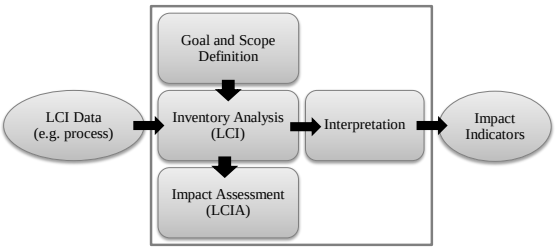


Figure 1. LCIs data adoption with four phases of LCAs procedure to yield an impact indicators

As illustrated in Fig. 1, the *inventory data (LCI data)* was collected in the format of standard guideline [10]. The LCI data is used in the four phases of LCA [11] that entail (1) explicating goals and scope, (2) elucidating and investigating a product inventory (Life Cycle Inventory {LCI}), (3) assessing an environmental impact (Life Cycle Impact Assessment {LCIA}), and (4) interpreting consequence. The last element is an *impact indicator*, which is an environmental impact with aspects of interest.

To represent LCA's domain knowledge, we investigate its data collection and assessment. Both of LCI and LCIA are considered and customized, while the other phases depend on user's particular specification; for example, goal and scope definition is explicated in the study and interpretation summarizes the results in report.

The structure of the LCA method as shown in Fig. 1 is used as the experiment guideline. The procedure requires the LCI data to be disseminated throughout the process flow. The flow model in LCI phase comprises data collection and data calculation involving the compilation and quantification of inputs and outputs. The LCI data contains relevant input and output flows for the whole life cycle of a product of interest. LCI phase is described as a flow chart comprising processes that are manipulated in the pertinent supply chain and draw the system boundary. To do so, all activities are defined by analyzing each process's input and output. In our study, there are two variants of LCI data involved *Cradle-to-Gate* (C2G: from resource extraction {cradle} to the factory gate), and *Gate-to-Gate* (G2G: considering only one value-added process

Table 1. Syntax and Semantics of Description Logic

Constructor name	Syntax	Semantics
Top concept	$\top$	$\Delta^{\mathcal{I}}$
Bottom concept	$\bot$	$\emptyset$
Conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
Disjunction	$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$
Negation	$\neg C$	$\Delta^{\mathcal{I}} \backslash C^{\mathcal{I}}$
Existential quantification	$\exists r.C$	$\{d \in \Delta^{\mathcal{I}} \mid \exists e : (d,e) \in r^{\mathcal{I}} \wedge e \in C^{\mathcal{I}}\}$
Universal quantification	$\forall r.C$	$\{d \in \Delta^{\mathcal{I}} \mid \forall e : (d,e) \in r^{\mathcal{I}} \rightarrow e \in C^{\mathcal{I}}\}$
Restriction (R minCardinality(n))	$\geqslant_{n.R}$	$\{a \in \triangle^{\mathcal{I}} \mid \lvert\{b \mid (a,b) \in R^{\mathcal{I}}\}\rvert \geq n\}$
Restriction (R maxCardinality(n))	$\leqslant_{n.R}$	$\{a \in \triangle^{\mathcal{I}} \mid \lvert\{b \mid (a,b) \in R^{\mathcal{I}}\}\rvert \leq n\}$
Concept inclusion	$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
Role inclusion	$r \circ s \sqsubseteq t$	$r^{\mathcal{I}} \circ s^{\mathcal{I}} \subseteq t^{\mathcal{I}}$
Concept assertion	$a : C$	$a^{\mathcal{I}} \in C^{\mathcal{I}}$
Role assertion	$a\, r\, b$	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$

in the entire production chain). For the next phase, the LCIA phase aims at understanding and evaluating the magnitude and significance of the potential environmental impacts of a product system. It is carried out on the basis of the LCI data. For the assessing process, there are three elements in the mandatory stage of LCIA involved impact categories, category indicators and characterization models. The assigned LCA data is classified (impact categories) and calculated to be the category indicator results (characterization).

In this paper, our intention is to minimize the assessed impact result. We applied the approach of Cleaner Technology [8; 12] which makes green buildings, transport, and infrastructure both more energy efficient and environmentally benign. There are three approaches of environmental impact (1) reduction includes material replacement, (2) process renovation, and (3) case-specific optimization. Among those, the replacement of input material is the most feasible approach to implement.

3 Framework Overview

3.1 Recommender System

To support implementation, the recommender system was divided into five layers: knowledge, data, operation, service, and application. Each tier of the recommender system is related to different DL elements. As shown in Fig. 2, both of knowledge and data layers are built and can be organized into five layers as follows: (1) **Knowledge Layer** (*TBox*): Source documents

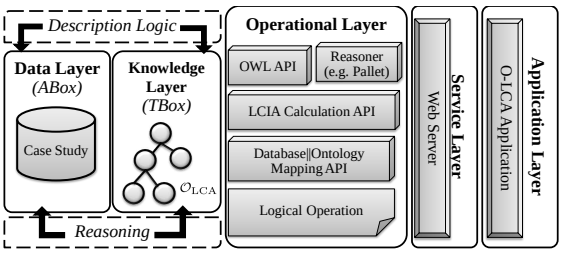


Figure 2. Description logics system with five layers of $\mathcal{O}_{\mathrm{LCA}}$ framework.

are gathered and expert suggestion is interpreted. The ontology structure entailing concepts and its relations is formalized as *TBox*. (2) **Data Layer** (*ABox*): Instances are collected and stored in the database as *ABox*. (3) **Operational Layer**: The next layer provides an application programming interface (API). (4) **Service Layer**: Web server software as Apache HTTP Server, provides service in database connection. We develop a web application following the Model-View-Controller Concept [13]. (5) **Application Layer**: The last layer representing the web application to client users with DHTML development [14].

3.2 Improvement of O-LCA Ontology

For the first O-LCA ontology ($\mathcal{O}_{\mathrm{LCA}}$), we designed it based on the related ontologies. Most of concepts to analyze are in the LCI Phase. ISO14048, data documentation guideline, was considered as the source document to collect terminological concepts in LCI phase. For example, the process involving its description, input

and output; modeling and validation, and administrative information are three initial parts of data document concepts. Each of those concepts requires a related concept involving substance, quantity, and its unit.

This phase, an impact assessment concept was analyzed for adding in a new $\mathcal{O}_{\mathrm{LCA}}$ that involves impact assessment method, aspect, factor value, and unit. By the course of adding new concept, the structure knowledge and relations will be changed and the related concepts are defined.

As shown in Fig. 3, TBox are separated into two parts. There are high-level concept of the new $\mathcal{O}_{\mathrm{LCA}}$as shown in Fig. 3(a). We initialize Process that comprises a variant of SystemBoundary and both input and output of QuantifiedFlow as shown in (a). The QuantifiedFlow (b) describes both flows with an existing substance (e.g., CrudeOil, Electricity), an interacting Environment, and the quantified elements with Quantity and Unit. The concept depicted in LCIA phase, ImpactAssessment (c) is defined based on ImpactAspect, its Unit that depend on the selected aspects (e.g., kg CO_2, kg SO_2) and related substance ImpactProperty. There are three defined elements in ImpactProperty (d): QuantifiedFlow, Unit of a selected impact, and value of impact factor. Process assessed by interested ImpactAssessment is an AssessedProcess (e) that quantify with the impact value.

The acquired case needs to be prepared and set as pre-process before implement. The ontology development guideline [15] is our approach. Starting with gathering information, we collect LCA resources such as source documents, and consult domain experts for the case (e.g. LCA's case of oil refinement). After that, with the ontology advantage in information sharing, the knowledge is described in usable and adaptable parts. There are two existed ontology that consist of CASCADE with data documentation ontology and OLCA with an excerpt part of ontology in the product flow. We analyze important terms in the source documents and previous work are enumerated. To make domain assumptions explicit, we analyze important terms in the source documents and the previous work and separate the domain knowledge from the operational knowledge. Those terms re-define the classes and their hierarchy, then properties and facets of each class are defined consequently.

Table 2. An LCA case study in oil refinery

INPUT		OUTPUT	
Substance	Quantity	Substance	Quantity
Crude oil	1.14 kg	Diesel	1 kg
Electricity	0.5 MJ	Carbon dioxide	0.226 kg
Chemicals	0.68 kg	Nitrogen oxides	0.024 kg
Fuel oil	0.259 kg	Methane	0.0513 kg
		Sulfer oxides	0.296 kg
		Benzene	0.00904 kg
		Xylene	0.00904 kg
		Toluene	0.0136 kg

However, the domain expert will be revised Ontology design and recognized LCA knowledge. The designed ontology will be consulted with domain expert in ontology comprehension and correctness. When we have the available ontology, the case study in oil refinery in Tab. 2 is created the instantiation. In context of the interested case, we have axioms in Fig. 3(b) to be asserted in Fig. 4. All related concepts with the specific case of oil refinery are explained in the next section.

4 Case Study

4.1 Oil Refinery

Our case study of interest with LCA knowledge domain is crude oil refinery. A refinery instruction [16] is sprawling industrial complexes with extensive piping running throughout, carrying streams of fluids between large chemical processing units. Following this instruction, the variations of the compounds boil at different temperatures and change to gases; and are later re-condensed back into liquids. The chemical process, in which crude oil is heated, is called fractional distillation. As shown in Tab. 2, the refinery process [1] requires four inputs distilled involving crude oil (or a petroleum), chemicals, electricity, and fuel oil. Outputs flow consists of oil products such as diesel and emissions involving Carbon dioxide (CO_2), Nitrogen oxides (NO), Methane (CH_4), Sulfur oxides (SO), Benzene (C_6H_6), Xylene (C_8H_{10}), and Toluene (C_7H_8).

As mention in the previous section, we have formalized the axiom of oil refinery case as

[1] Motivated by personal communication with an MTEC (National Metal and Materials Technology Center) researcher.

(a)	Process	$\sqsubseteq$	$\exists$hasVariant.SystemBoundary $\sqcap$ $\exists$hasInputFlow.QuantifiedFlow $\sqcap$ $\exists$hasOutputFlow.QuantifiedFlow
(b)	QuantifiedFlow	$\equiv$	Flow $\sqcap$ $\exists$hasSubstance.Substance $\sqcap$ $\exists$interactWith.Environment $\sqcap$ $\exists$hasQuantity.Double $\sqcap$ $\exists$hasUnit.Unit
(c)	ImpactAssessment	$\sqsubseteq$	$\exists$hasImpactAspect.ImpactAspect $\sqcap$ $\exists$hasImpactUnit.Unit $\sqcap$ $\exists$hasImpactPpt.ImpactProperty
(d)	ImpactProperty	$\sqsubseteq$	$\exists$hasImpactFlow.QuantifiedFlow $\sqcap$ $\exists$hasFactor.Double $\sqcap$ $\exists$hasFactorUnit.Unit
(e)	AssessedProcess	$\sqsubseteq$	$\exists$selectedCase.Process $\sqcap$ $\exists$assessedBy.ImpactAssessment $\sqcap$ $\exists$hasImpactValue.Double

(a) TBox: Top concepts

$(f1)$	OilRefinery	$\equiv$	Process $\sqcap$ $\exists$hasVariant.G2G $\sqcap$ $\exists$hasInputFlow.QFlow_CrudeOill $\sqcap$ $\exists$hasInputFlow.QFlow_Electricity ... $\sqcap$ $\exists$hasOutputFlow.QFlow_Diesel $\sqcap$ $\exists$hasOutputFlow.QFlow_CarbonDioxide
$(f2)$	EG_Coal	$\equiv$	Process $\sqcap$ $\exists$hasVariant.C2G $\sqcap$ $\exists$hasInputFlow.QFlow_CoalEnergy ... $\sqcap$ $\exists$hasOutputFlow.QFlow_Electricity
$(f3)$	EG_Wind	$\equiv$	Process $\sqcap$ $\exists$hasVariant.C2G $\sqcap$ $\exists$hasInputFlow.QFlow_WindEnergy ... $\sqcap$ $\exists$hasOutputFlow.QFlow_Electricity
$(f4)$	EG_Hydro	$\equiv$	Process $\sqcap$ $\exists$hasVariant.C2G $\sqcap$ $\exists$hasInputFlow.QFlow_WaterEnergy ... $\sqcap$ $\exists$hasOutputFlow.QFlow_Electricity
$(g1)$	QFlow_CrudeOil	$\equiv$	QuantifiedFlow $\sqcap$ $\exists$hasSubstance.CrudeOil $\sqcap$ $\exists$interractWith.Technosphere $\sqcap$ $\exists$hasQuantity.Double $\sqcap$ $\exists$hasUnit.Unit_kg
$(g2)$	QFlow_Electricity	$\equiv$	QuantifiedFlow $\sqcap$ $\exists$hasSubstance.Electricity $\sqcap$ $\exists$interractWith.Technosphere $\sqcap$ $\exists$hasQuantity.Double $\sqcap$ $\exists$hasUnit.Unit_MJ
		$\vdots$	
$(g3)$	QFlow_Diesel	$\equiv$	QuantifiedFlow $\sqcap$ $\exists$hasSubstance.Diesel $\sqcap$ $\exists$interractWith.Technosphere $\sqcap$ $\exists$hasQuantity.Double $\sqcap$ $\exists$hasUnit.Unit_kg
$(g4)$	QFlow_CarbonDioxide	$\equiv$	QuantifiedFlow $\sqcap$ $\exists$hasSubstance.CarbonDioxide $\sqcap$ $\exists$interractWith.Technosphere $\sqcap$ $\exists$hasQuantity.Double $\sqcap$ $\exists$hasUnit.Unit_kg
(h)	GreenHouseEffect	$\equiv$	ImpactAssessment$\exists$hasImpactAspect.GreenHouse $\sqcap$ $\exists$hasImpactUnit.Unit_kgCO2 $\sqcap$ $\exists$hasImpactPpt.ImpactProperty
(i)	ImpactProperty	$\sqsubseteq$	$\exists$hasImpactFlow.CarbonDioxide $\sqcap$ $\exists$hasFactor. "1" $\sqcap$ $\exists$hasFactorUnit.Unit_kgCO2_perkg $\sqcap$ $\exists$hasImpactFlow.Methene $\sqcap$ $\exists$hasFactor. "11" $\sqcap$ $\exists$hasFactorUnit.Unit_kgCO2_perkg
(j)	CleanProcess_Case1	$\equiv$	AssessedProcess $\sqcap$ $\exists$assessedImpact.GreenhouseEffect $\sqcap$ $\exists$hasImpactValue.$(\leqslant_{1.7})$

(b) TBox: Specific case study

Figure 3. A core extract of the improved $\mathcal{O}_{\mathrm{LCA}}$

shown in Fig. 3(b) from $(f1)$ to (j). The process initialize with OilRefinery $(f1)$ concept and it description involves Gate-to-Gate (G2G) as the SystemBoundary and both input and output flows. Note that, for $(f1)$, there are many QuantifiedFlows as show in Table. 2. Our example flows will represent only two of input flows (QFlow_CrudeOil $(g1)$ and QFlow_Electricity $(g2)$) and two of output flows (QFlow_Diesel $(g3)$ and QFlow_CarbonDioxide $(g4)$).

To prepare an alternative energy, there are three sources of energy for *electricity generator (EG)* that consisted of EG_Coal $(f2)$ (an electricity generator from coal), EG_Wind $(f3)$ (an electricity generator from wind turbine) and EG_Hydro $(f4)$ (an electricity generator from water). Those processes have variant as Cradle-to-Gate (C2G). Although their provenances is difference, they production is QFlow_Electricity that is used as the input flow of OilRefinery. Note that, those alternative processes will be represented only important input flow and utilized output flow.

For process assessment, an impact method with environmental aspect was described in GreenHouseEffect (h). The concept inherited from ImpactAssessment (c) and responded with green house aspect, impact Unit and ImpactProperty (i) that list all substances and impact factors with their weight and unit. To follow the Cleaner Technology approach, the validity of CleanProcess (j) is defined by AssessedProcess which is specific impact aspect and not more than specific impact value.

In the context of oil refinery, axioms in Fig. 3(b) augmented with assertions in Fig. 4 form LCA knowledge base. An instance of OilRefinery concept is `OilRefinery01` at $(a1)$, and three alternative input flows are instantiated in $(a2)$, $(a3)$ and $(a4)$ in Fig.4. To fulfill the flows, input flows are asserted as the role of `OilRefinery01` involves in $(b1)$, $(b2)$, $(b3)$ and $(b4)$. Note that those flows are represented only two of each input and output flows. With the role assertion for $(a1)$, the flows are quantified by instantiation as follows $(c1)$, $(c2)$, $(c3)$ and $(c4)$. After constructed the instances the oil refinery and three alternative energy resource, they are calculated the environmental impact. Following the cleaner technology's approach to checking CleanProcess_Case1 as defined in (j) of Fig. 3(a), an assessed process as an AssessedProcess instance is asserted as

$(a1)$	`OilRefinery01`	: OilRefinery
$(a2)$	`EG01_Coal`	: ElectricGenerator_Coal
$(a3)$	`EG02_Wind`	: ElectricGenerator_Wind
$(a4)$	`EG03_Hydro`	: ElectricGenerator_Hydro
$(b1)$	`OilRefinery01`	hasInputFlow `CrudeOil01`
$(b2)$	`OilRefinery01`	hasInputFlow `Electricity01`
		⋮
$(b3)$	`OilRefinery01`	hasOutputFlow `Diesel01`
$(b4)$	`OilRefinery01`	hasOutputFlow `CO2_01`
$(c1)$	`CrudeOil01`	: QFlow_CrudeOil ⊓ ∃hasQuantity. "1.14"
$(c2)$	`Electricity01`	: QFlow_Electricity ⊓ ∃hasQuantity. "0.5"
		⋮
$(c3)$	`Diesel01`	: QFlow_Diesel ⊓ ∃hasQuantity. "0.226"
$(c4)$	`CO2_01`	: QFlow_CarbonDioxide ⊓ ∃hasQuantity. "0.0136"
$(d1)$	`AsPrc_OR_01`	: AssessedProcess
$(d2)$	`AsPrc_OR_01`	selectedCase `OilRefinery_01`
$(d3)$	`AsPrc_OR_01`	assessedBy GreenHouseEffect
$(d4)$	`AsPrc_OR_01`	hasImpactValue. "1.8585"
$(d5)$	`AsPrc_OR_02`	hasImpactValue. "1.6585"
$(d6)$	`AsPrc_OR_03`	hasImpactValue. "1.7085"

Figure 4. An ontological assertion describing logical structure of $\mathcal{O}_{\mathrm{LCA}}$ that describes with DLs in case of oil refinery.

`AsPrc_OR_01` in $(d1)$. The instance is verified with selected case study $(d2)$ as `OilRefinery01` and assessed by GreenHouseEffect in $(d3)$ as defined in (h) of Fig. 3(b). There are three impact AssessedProcesses as follows:

- The assessed process $(d4)$ by electricity from coal $(a2)$ has the environmental impact 1.8585 kg CO_2.
- The assessed process $(d5)$ by electricity from wind $(a3)$ has the environmental impact 1.6585 kg CO_2.
- The assessed process $(d6)$ by electricity from water $(a4)$ has the environmental impact 1.7085 kg CO_2.

With the given reasoning, we can infer that the oil refinery process with using from wind is the clean process, because the environmental impact is lower than 1.6 kg CO_2 as defined by (j) concept in Fig. 3(b).

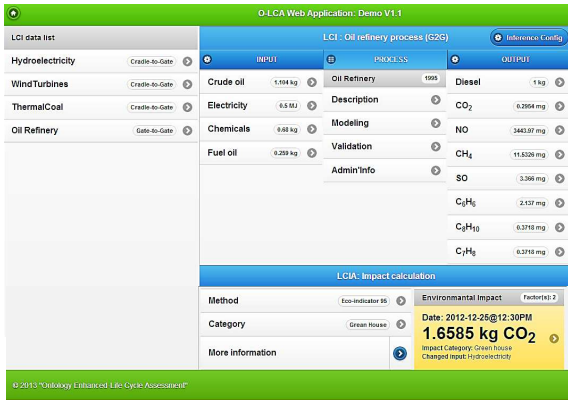


Figure 5. An interface of $\mathcal{O}_{\mathrm{LCA}}$web application

For implementation of $\mathcal{O}_{\mathrm{LCA}}$ application, Graphic User Interface (GUI), we design a web application interface. The application layout is separated into three parts list of case study at the left, LCI phase at the top right, and LCIA bottom right. However, to make it user friendly, jQueryUI for mobile was chosen to attach user, and can be used in mobile device as shown in Fig. 5.

5 Evaluation and Discussion

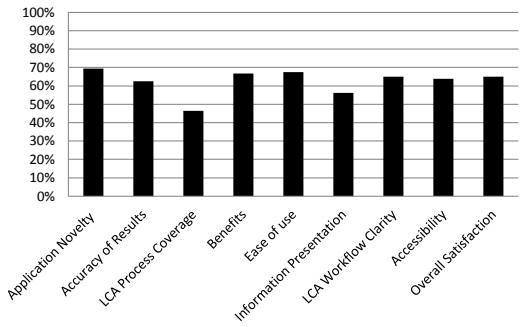


Figure 6. System evaluation result

$\mathcal{O}_{\mathrm{LCA}}$ application has been evaluated by means of a questionnaire. The weighted average of satisfaction aspects is calculated. The granularity of satisfaction is: strongly agree (4 points), agree (3), neither agree or disagree (2), disagree (1) and strongly disagree (0). The total weighted score in each approach is then divided by the total weighted score. As shown in Fig. 6, they are presented as the percentages of the weighted average score of each aspect. We created and distributed a survey focusing on user satisfaction towards our system. The total number of participants is 10 people; consisting of 6 LCA experts and 4 ontology engineers. None of respondents possess both LCA and ontology development as background knowledge.

The evaluation routine was to let all the respondents use application to calculate environmental impact and recommend an alternative input of the process flow, and to let them complete the survey questions regarding their satisfaction. Then, it let them suggest improvement on coverage of LCA knowledge and the ontological structure. We provide both calculated environmental impact and recommended an alternative input. After that, the respondents were asked to complete the survey.

The results in Fig. 6 show that most of the aspect more than 60%. Application novelty has the highest score while ease of use and benefit are also highly satisfied. However, system coverage is still unsatisfied by less than 50% of the participants. We classified the user feedback and suggestions for the system into three points. First, there is no clear definition of *absolute cleanness* in practice; only relative cleanness is considered. Our system should record the least impact result and use it as the criterion of cleanness. Second, our system was designed only for single-producted process. In practice, multiple-producted processes require impact allocation with respect to the product ratio. Last, the impact representation should be dynamic to the change of rules.

6 Conclusion

We have presented a recommending system for life cycle assessment (LCA) by incorporating the notion of the ontology in formalizing its procedure. The formalized knowledge structure is based in DL or $\mathcal{O}_{\mathrm{LCA}}$ was implemented with the knowledge application framework. The application was applied to suggest the material replacement approach with the case study of oil refinery. The $\mathcal{O}_{\mathrm{LCA}}$ performance was utilizing by express assessment procedure involved LCI data, LCI and LCIA with DLs syntax.

However, to improve the performance of $\mathcal{O}_{\mathrm{LCA}}$, we plan to solve some issues. With the lack of context interpretation, we will find out the technique to change the process status as role, e.g. Semantic Markup for Web Services (OWL-S) will be explore with the performance in functionalities of descriptive language. And another issue such as the uncontrollable factors

(e.g. valid localization, year), we plan to use the Data Quality Indicator (DQI) to enhance the performance of data reliability which will help LCA in decision making in data selection.

Acknowledgment

This research is financially supported by Thailand Advanced Institute of Science and Technology (TAIST), National Science and Technology Development Agency (NSTDA), Tokyo Institute of Technology, Sirindhorn International Institute of Techcnology (SIIT), Bangchak Petroleum Public Company Limited, Thammasat University (TU) and the National Research University Project, Thailand Office of Higher Education Commission. Materials and information provided by National Metal and Materials Technology Center (MTEC, Life Cycle Assessment {LCA} Laboratory, Thailand). The authors are grateful to Marut Buranarach and Prachya Boonkwan for fruitful technical discussions.

References

[1] G. Schreiber and M. Dean. OWL web ontology language reference. W3C recommendation, W3C, February 2004.

[2] A. Takhom et al. Ontology-enhanced life cycle assessment: Toward formalizing the standard guidelines. In *The International Conference on Information and Communication Technology for Embedded Systems*, 2013.

[3] I. Horrocks F. Baader and U. Sattler. Description Logics. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, chapter 3, pages 135–180. Elsevier, 2008.

[4] B. Suntisrivaraporn. *Polynomial-Time Reasoning Support for Design and Maintenance of Large-Scale Biomedical Ontologies*. PhD in Computer science, Technischen Universitaet Dresden – Fakultaet Informatik, 2009.

[5] M. Braescher et al. Cascade. In *The 16th Internationale Conference: informatics for environment protection*, 2002.

[6] Marisa Braescher, Fernand Monteiro, and Alessandra Silva. Life cycle assessment ontology. Congreso ISKO-España, 2007.

[7] B. Bertin et al. A semantic approach to life cycle assessment applied on energy environmental impact data management. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops*, EDBT-ICDT '12, pages 87–94, New York, NY, USA, 2012. ACM.

[8] R. Jain and S. K. Sikdar. *Clean Technologies and Environmental Policy*. Springer Verlag, 2004.

[9] *Life Cycle Assessment (LCA) - A guide to approaches, experiences and information sources*. EEA (European Environment Agency), 2007.

[10] Organización Internacional de Normalización. *ISO TS 14048: Environmental Management, Life Cycle Assessment, Data Documentation Format*. ISO, 2002.

[11] International Organization for Standardization. *ISO 14044: Environmental Management - Life Cycle Assessment - Requirements and Guidelines*. ISO, 2006.

[12] M. Getzner. The quantitative and qualitative impacts of clean technologies on employment. *Journal of Cleaner Production*, 10(4):305 – 319, 2002.

[13] G. Krasner and S. Pope. A description of the Model-View-Controller user interface paradigm in the smalltalk-80 system. *Journal of Object Oriented Programming*, 1(3):26–49, 1988.

[14] C. Powell and K. Akama. Structured development of dhtml programs from abstract ideas based on the equivalent transformation framework. In *Innovative Computing, Information and Control, 2007. ICICIC '07. Second International Conference on*, page 384, sept. 2007.

[15] N. F. Noy and D. L. Mcguinness. *Ontology Development 101: A Guide to Creating Your First Ontology*. 2001.

[16] H. Lee M. Wang and J. Molburg. *Allocation of Energy Use in Petroleum Refineries to Petroleum Products: Implications for Life-Cycle Energy Use and Emission Inventory of Petroleum Transportation Fuels*, 9(1):34–44, July 2003.

A Content Management System for Food Industry

Tomoko Kashima † **Shimpei Matsumoto** ‡
†Faculty of Engineering, Kinki University
‡Faculty of Applied Information Science, Hiroshima Institute of Technology
†kashima@hiro.kindai.ac.jp
‡s.matsumoto.gk@cc.it-hiroshima.ac.jp

Abstract

There are many problems in Japanese agriculture. To overcome the dwindling and graying population of farmers, the authors have developed an information sharing system of food production, marketing, and consumption to promote its proper management for food industry, named food CMS. The system provides services appropriate to the types of users, farmers, customers, and distributors including retailing stores, brokers, and restaurants. Each user can reap benefits according to one's role from participating in this system, not only the direct relationship-building between customers and producers which has long been difficult, but also freshness management of foods, efficiency and rationalization of ordering works in distributors, and expansion of sales opportunity. All kinds of food distribution information are stored in the integrated database, so the information can realize to adjust supply to demand. The system will contribute to save the impact on the environment because it can reduce the food loss by promoting production with just needed kinds and needed quantity when they are needed. This paper shows the effort the authors have addressed, especially the function design of food CMS.

Keywords: Farmers Information System, Agribusiness, Farming, Demand Forecasting

1 General Instructions

Recently, many farmers are experiencing a lot of problems. Farmers are rapidly aging because young people who get an agricultural job are decreasing in number. The average age of agricultural workers in Japan is over 66 years. The percentage of the population engaged in agriculture in Japan has been about 2% of the workforce since 2002, falling from 26.8% in 1960. The percentage of GDP (Gross Domestic Product) accounted for by agricultural production also dropped from 9% to less than 1% during the same period. The importance of agriculture to the Japanese economy has rapidly diminished in Japan with economic growth. Everyone understands the necessity of agricultural reform. The importance of agriculture and food supply issues take time to improve the agriculture industry. But further structural reform necessitates radical reform of agricultural fields.

Over the years a number of laboratories have studied optimization of farm work using a sensor, camera and devices [1][2]. Researchers have recently proposed the development of hardware to support the farmers. But the literature is concerning system to help farmers by providing information. However, their hardware is expensive and limited. New business models for agricultural markets have needed using IT.

The authors have already developed a management function in Farmers' Information System (FIS) to obtain skillful farmers' knowledge [3]. The authors will have a great store of experience of farmers from collecting information using FIS to support and cultivate young producers, who would shoulder the task of upholding agriculture in the future. And, the farmer was able to encourage value creation of agricultural products using FIS by giving a customer the obsessiveness to each one of products, and additional information directly. From the operation experiment of single system, it was checked that it is effective in the support which builds the relation between consumers and farmers.

In this research, we pay attention to information sharing of the point of production which is a starting point of circulation, and the point of consumption which is the last points. We develop the information sharing system related to production, circulation, and consumption. We dig up the request from the consumer side for the improvement in manufacturing efficiency, and think that it is necessary to assign farm prod-

ucts exactly according to the feature. We think that they become possible by development of the information system based on the feature of SNS. The proposed system, named Food CMS, is developed with the previous efforts of FIS. A system offers a needed function according to a user's role. By this system, it is possible that to build relationships with the conventionally difficult customer and a producer, preserve freshness of food, improve the efficiency in sales order of distribution center's buyer and expansion of sales opportunities. The user can get the profits according to each role.

2 Related Work

Previous agricultural research was aimed at computerization. Support of the traceability aiming at branding of agricultural products, optimization of the harvest which aimed at higher levels of productivity, make use of sensing technology and acquisition of the expert knowledge by an agricultural-work history. For example, the paper by takatsu et al. reported that there is introduction of ICT as a solution of agricultural successor problem. Additionally, the introduction of ICT contributes to know of the harvest time of agricultural products and prediction of yield amount and to improve of farmers income [4]. In order to put introduction of ICT to practical use, it is necessary to reduce the cost of providing a service. Therefore, the authors say that practical use of Cloud service is effective. In order to put introduction of ICT to practical use, it is necessary to reduce the cost of providing a service. Therefore, the authors say that practical use of Cloud service is effective. Similarly, Saito et al. also described the effectiveness of information system construction [5]. Saito et al. state that information of safety and security of food can contribute to branding of agricultural products. Then Mechanism of integrated management of distribution information is needed. So The authors start to develop Food CMS.

There is keen interest to safety and security of food and agricultural environment protection in Japanese society. There are a lot of technologies which is focus attention on this point. However, interest of many researchers are focused on limited area of society. So, Technology does not receive and spread easily in society in Japan. For example, Hayashi et al. has pointed out some problems which have one risk modification raise another issue, or trade-off between risk modification and primary advantage [6]. Nanseki et al. proposed the next generation food system which can supply food continuously by the same point of observation [7]. The problem consciousness shown by precedence study is the study background of total foods market information control package development.

3 Food CMS

3.1 Concept

Previously the authors have developed Farmers Information System (FIS), and FIS is new business models for agricultural markets have needed using IT[3]. The problem of food attracted attention since about 2000 in Japan. Consumers changed to become interested to food of secure and safe. Consumers began to ask for getting to know how and where agricultural products were made. Consumers came to get interested also in quantity of the agricultural chemicals of agricultural products and growing information of agricultural products. Because of this situation, the purpose of this FIS is "Change the domestic agriculture using IT".

1. Management stability of agriculture.
2. Stable supply of agricultural products.
3. Improvement in the value of agricultural products.

Under this perspective, the authors develop a new information system named Food CMS for urban markets to facilitate the transactions. Food CMS is CMS having a marketing function and SNS, and is the system which specialized in agricultural product marketing. Traditional, as a circulation package of goods, it is EC Cube,and as a SNS package, it is OpenPNE are developed and these most are exhibited by the open source. However, it has both feature and there is no package for exclusive use which specialized in food circulation further. In the case of agricultural products, there is an element of raising products, and especially the traceability of the growth process is searched for. Therefore, it differs from an industrial commodity or daily necessaries. And, because agricultural products are living things, preservation and management are difficult for them. Then, consumers' voice is sent to Farmers, farmers make it possible to produce agricultural products according to consumers' request

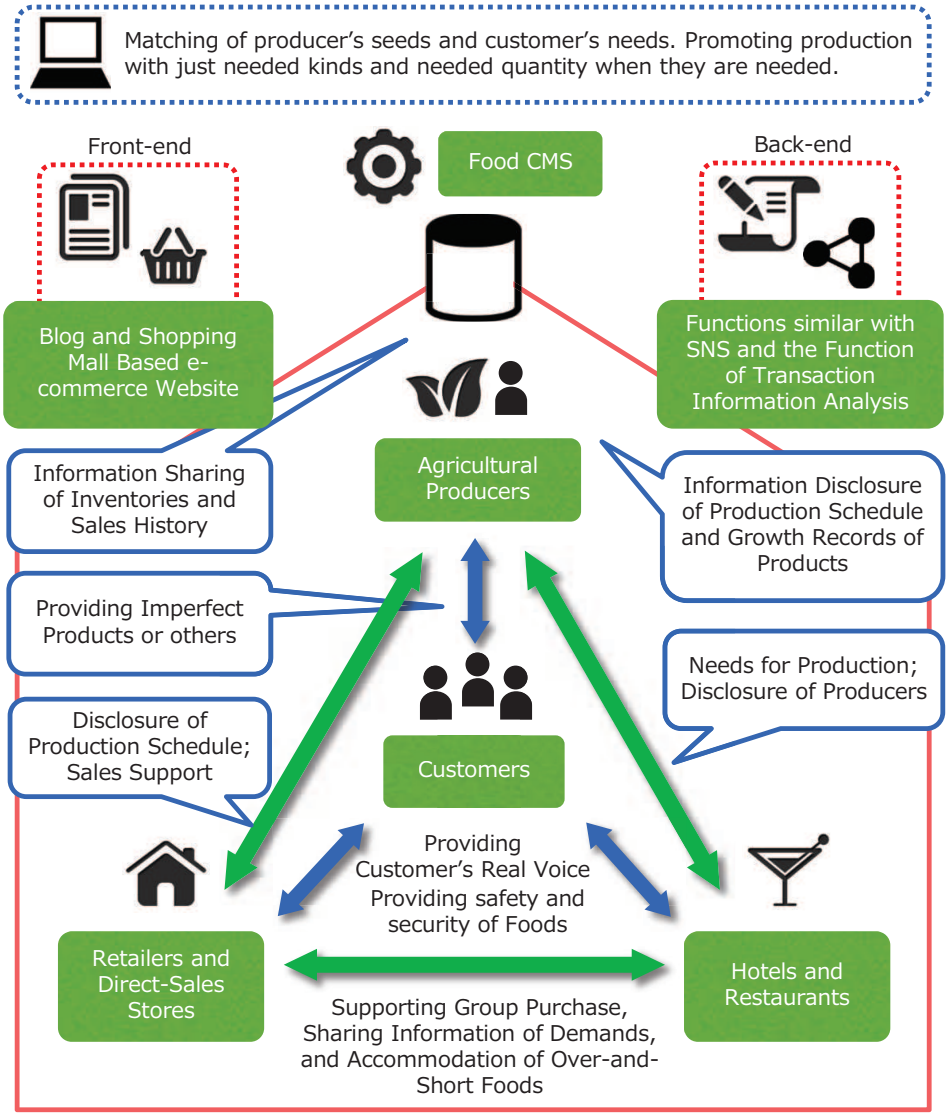


Figure 1. Concept of Food CMS

only a suitable quantity for suitable time. In order to satisfy the demands of above request, the feature of SNS is very effective. For the background of having resulted in authors' idea. There is already a report which expected cooperation of a restaurant and an agricultural-products direct sales store [8]. There is a report which described the shared necessity of the consumption trend of farm products, production information, and physical distribution information [9]. It is thought that an important issue is cooperation with the producer of a direct sales store or a restaurant group especially.

3.2 Function of Demand Forecasting

Food CMS consists of two services, a front end and a back end. In the front end, it has a user interface of Web page form, and the back end has an interface of the setting screen of CMS, and a function of SNS. All the users of a system are provided with the function of a front end and a back end. The class structure shown in Figure 2. This is a front end of Farmers and is an example at the time of applying a default skin as it was and utilizing. And, Satake et. al did sharing of a consumption trend, production information, and physical distribution information. And wasteful reduction which avoids abandonment of products Improvement in the freshness

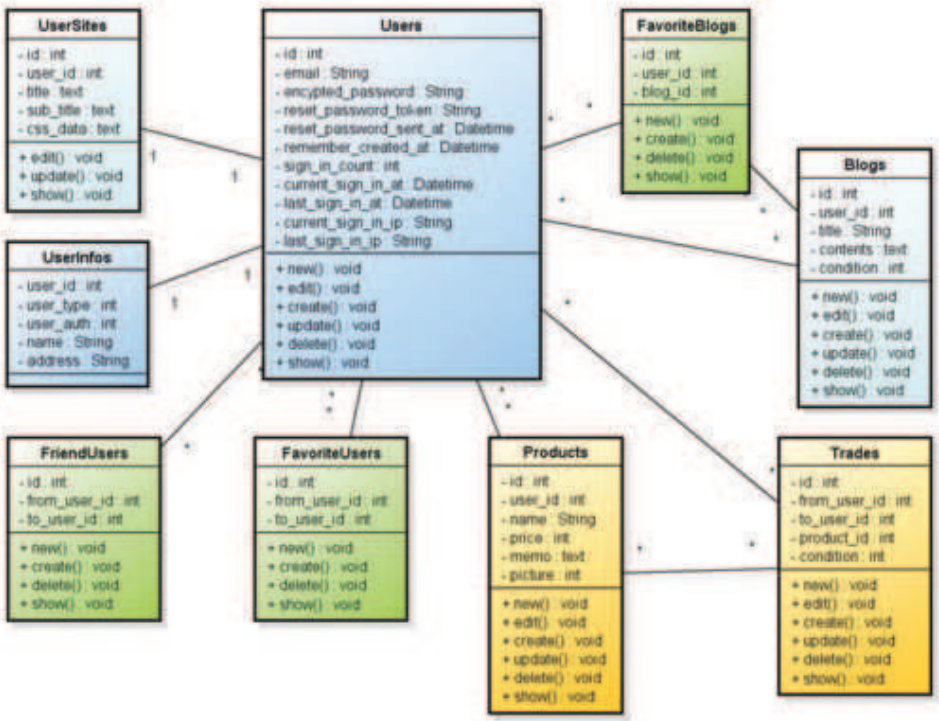


Figure 2. Internal Structure

of the products to ship Value, such as production / processing information, is raised [9]. The view of these Satake et.al agrees with authors' awareness of the issues. Not only as for bidirectional information dissemination simply but the taste according to consumers' attribute becomes clear through the accumulation of a buying history, the feature of SNS is added to the e-commerce function for dealings is expected. Matching of producer seeds and consumer needs is attained as the result.

3.3 Details of Front and Back Service

The general public who is that user is not registered can also browse the front end of the system. Browse the system authority is not granted in all the pages of a system, but Browse the system authority is set up for a portion of a page. The design layout of a front end can be performed for user itself. It is the form which uploads a CSS file. The design of two or more patterns is prepared by the default.

The Food CMS would include pictures of the actual farmers and their farms, information on growing practices and methods will also be included, along with the information on what type of chemicals or fertilizers were used in growing the products, or what type of preservatives were added to the products, if ever there was any. The objective of the system is to provide consumers information that cannot be known by simply looking at the products. This information can be considered as a value added to the products, thereby resulting to better acceptance of the products and better market prices. Figure 3 and Figure 4 has useful information which posted message and photos from farmers, consumers, seller and sommelier of vegetables. Uses of the system don't need active use, but need for only sending an e-mail with a photo attached. So this system is easy to use.

In a Figure 6, a back end is shown. It is provided only for the user by whom the user registration is done. This figure is a management screen of Farmers. Screen composition is common to all the users. However, the composition of a left-hand side menu and its number change with a user's rolls. A system construction function is mainly offered. For example, a CMS function, EC function, a SNS function, and the statistics analysis feature of a history are offered. Other addition functions, for example, a planning-of-production function, are additionally offered according to a user's roll. Three rolls

Figure 3. Website of a Direct-sales Place

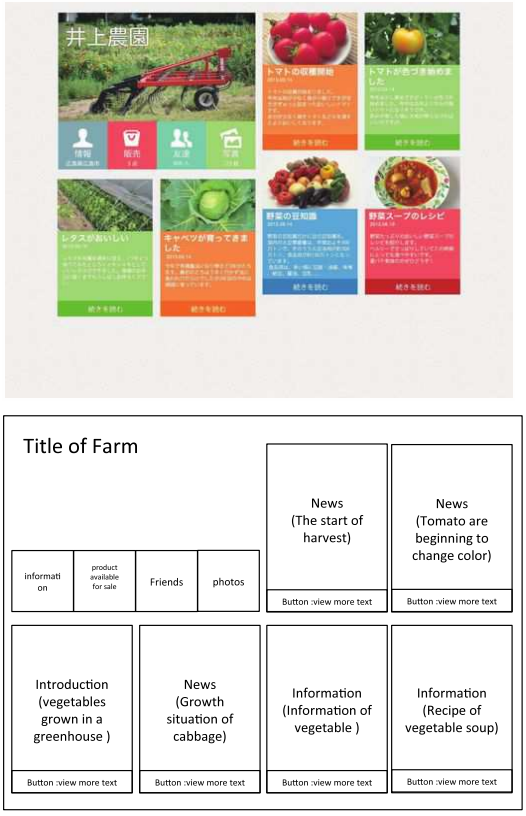


Figure 4. Information transmission utilizing IT

are prepared for the user. They are a producer, a distributor, and consumers. The user can serve as two or more rolls. The class shown in the figure is realized (see Figure 5 and 7).

4 Result and Effect of the Introduction

The authors were introduced to urban farmer's market in Hyogo in Japan the developed system. In this section shows effect of introduced. This time, the raw data such as sales cannot be displayed in this paper by contract.

First, the authors report the results of analysis of the system about the visitor. Figure 6 shows the percentage of visitors and new visitors of the system. Visitors are done by using a cookie to identify a browser on a particular computer as a single visitor. The most accurate visitor-tracking systems generally employ cookies to maintain tallies of distinct visitors. There are two types of visitors. New visitors are the first time a browser accesses the system. The system records the visitor as new. This is done by checking to see if the urchin utm cookie for our domain exists on the browser. If it does not, the visit is con-

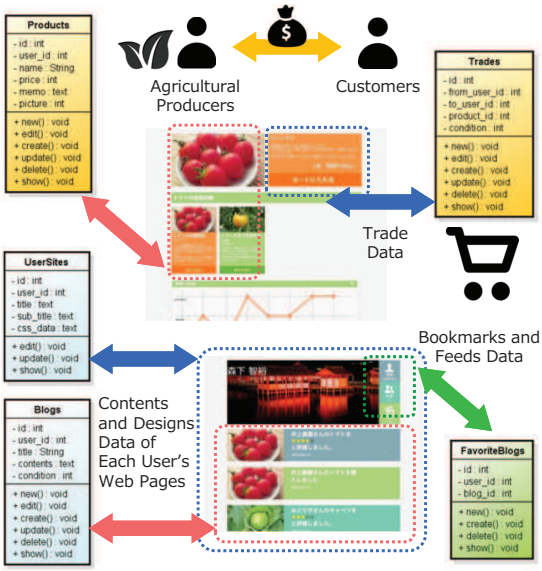


Figure 5. Internal Design of Front-End

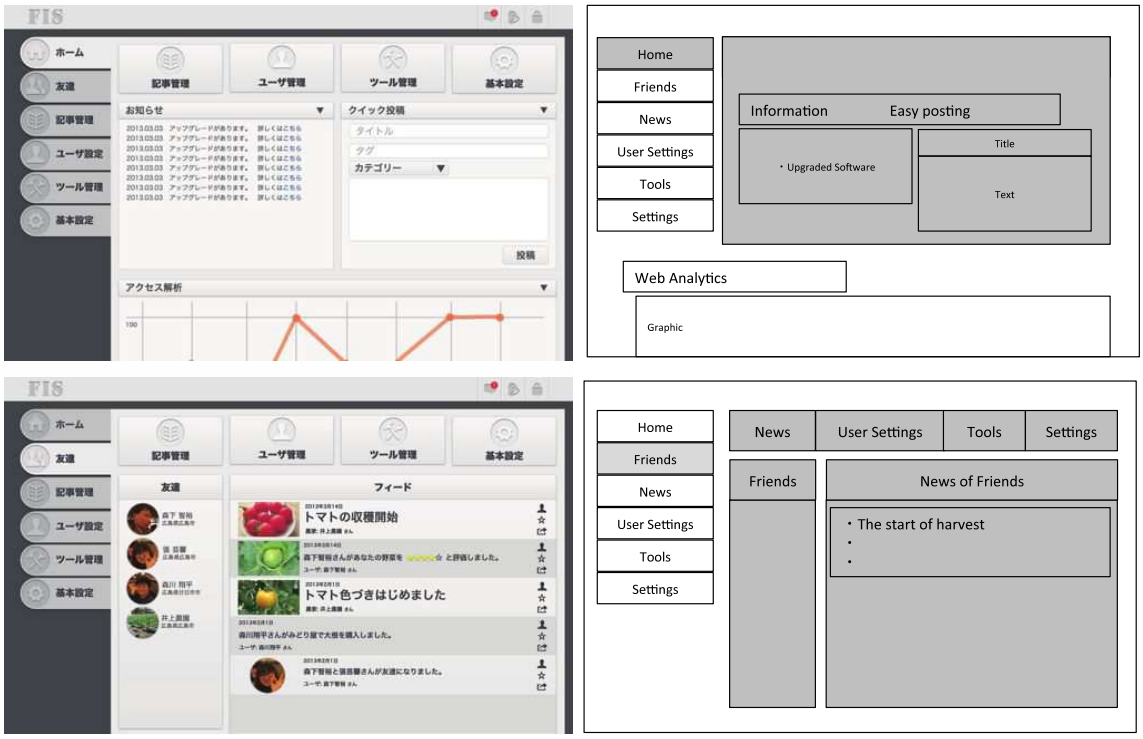


Figure 6. SNS function of back-end services

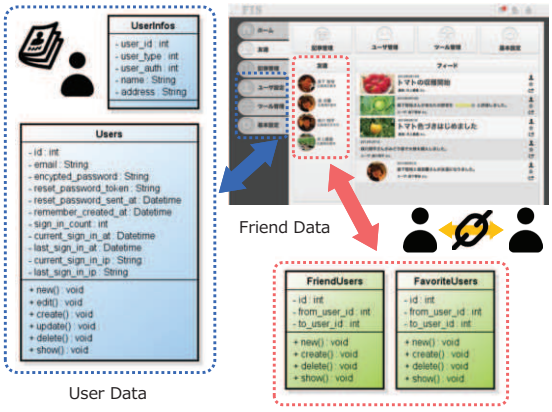


Figure 7. Internal Design of Back-End

sidered a first-time visit. Returning visitors is the system records a visitor as returning when the urchin utm cookie for our domain exists on the browser accessing the system. The authors got the data of the number of visitor to a system immediately after introduction of a FIS system. In this period, SEO and listing advertising measures have not done. New visitors accounted for approximately 60% in spite of not doing the advertising. These results show that FIS was viewed online by customers who wants to know safety and security of food. And, number of returning visitors have been gradually increasing from the introduction stage. The increase of returning visitors is considered that the user was interested in the information on a site. FIS believed that shares great amounts of information between farmaer and customers. Other page lets you see where visits originate. Location is derived from mapping IP addresses to geographic locations. High percentage of Hyogo Prefecture can be predicted. Because, shop of the experiment, in the Hyogo Prefecture. However, visitors have about 17% of visitors from Tokyo(over 500 kilometers away). The shop is acquired the user outside the prefecture by introduction of the system. Users from outside of the prefecture are

browsed information of the system and are felt safety and security of agricultural food. Thus, the shop has increased the demand for mail order from Tokyo. Sales of the shop has become a 140% year-on-year due to these factors. The authors resolve the agricultural problem through the system in this way. But, some agricultural crops don't have the opportunity to harvest only once a year. So to show the further effect that require more time than three years.

	B to B Commerce	B to C Commerce	Social Functions	Analysis Functions	Added Value
Telefarm		✓	✓		Game and local contribution
Radishbo-ya		✓			Safety and security
Agrisaurus (Farmscape)				✓	Knowledge sharing
Farmers WEB Farmigo	✓				Local production for local consumption
Good eggs		✓			Local production for local consumption
Food CMS	✓	✓	✓	✓	Insentive for customers

Figure 8. Compared to Other Similar Service

5 Discussion

Both sides, consumers and farmers, require certain information from markets about agricultural products. For example, this system are used by two different types of user. One is the type who is the actively write and utilize the function of food CMS. For example, it is agricultural producers. They will be posting information disclosure of production schedule and growth records of products and providing safety and security of food. Another type is read-only user. For example, it is hotels and restaurants. They will not be actively posting. But, the photo posted on system can show restaurant's customers information. In fact, Consumers may make requests about the exact information of agricultural products or their safety, while farmers may want to boast with their products. FIS can assist in the negotiation among both parties. If the farmers do not adjust their farming schedule to fit the demand of consumers, they automatically loose that sales opportunity. It is then very important that the farmers take into consideration what the consumers really want. The authors will avoid that their products are sold at a very low price. Farmers should focus on producing products that has actual consumer demand. FIS can adjust their farming schedule in a way that by the time they grow the product, their product is actually sought upon by consumers.

Because of these efforts, the authors provided infinite opportunities that farmers disclose, and created a value-creation opportunity for safety of agricultural crops. For example, the system disclose information which spray crops with pesticides, record growth of agricultural crops and serving idea. The system feeds back the result of basic function and have a functional evolution at farmer's request. The authors summarized development results as a Content Management System(CMS) package of the integrative information sharing system which the persons involved in food production, circulation, and consumption can use and aim at public presentation by an open source.

This system and similar service have been already in operation. The contribution to local supply and local consumption or farm management, e-commerce, the information-sharing function of a producer and consumers, etc. are implemented. A Figure8 shows what summarized those functions. A Telefarm and a Radishbo-ya are services of Japan. Others are U.S. service. A Telefarm does remote cultivation of the organically grown vegetable on the Internet using a personal computer at home, a personal digital assistant, a game terminal, etc. The Telefarm aims at activation of a rural area, and a contribution in the safe side of food. The Radishbo-ya sells organic matter and low-pesticide vegetables, additive-free food, etc. Because it is a major company, there are much number of articles and information and the reliability of information is high. Farmscape is service supported so that ordinary persons can also enjoy a kitchen garden freely. If it applies for service, the planter which cultivates vegetables will arrive and Farmers will come for help regularly. "urban-type of agriculture" is a keyword and is introduced into Los Angeles at 300 or more places. Agrisaurus has some feature that registers vegetables, inform you to watering schedule, up to the timing of harvest. Agrisaurus is a tool which can manage various things. In Japan, "Cropnet", "vegetable garden Navi", etc. which are SNS for a kitchen garden attract attention. Food CMS is service of a free participatory type. It is thought that technical solution is needed in the field of the reliability of information.

6 Conclusion

In this research, the problem about the present agriculture was shown. Farmers information system for carrying out a remedy to three problems was proposed. Solution was proposed to those problems and the support system was developed. The authors developed a simple prototype of FIS, the database of the proposed system will be examined to further cooperate with the stabilization of agricultural management and the support for stable supply of agricultural products. The structure that connects farmers, consumers, and a store was built using FIS. The major merit cooperating with agricultural field is to ensure traceability of foods for customers, and it lead to improve the service quality. And there will be new knowledge with the data collected by FIS.

Acknowledgements

The authors are very grateful to member of Institute of Agricultural Products Distribution Research, for their valuable cooperation in experiments. This work was spported by Japan Society for the Promotion of Science, Grant-in-Aid for Young Scientists(B), No.25750007 and A-step feasible study stage, No. AS242Z02070H.

References

[1] T. Togami, R. Ito, A. Hashimoto and T. Kameoka, Agro-Environmental Monitoring Using A Wireless Sensor Network to Support Production of High Quality Mandarin Oranges, *Japanese Society of Agricultural Informatics*, Vol.20, No.3, pp.110–121(2011).

[2] K. Kamimura, The management problem of Japanese agriculture, *Journa of Management consulting*, Vol.55, No.9, pp.28–31(2008).

[3] T Kashima, S Matsumoto and T Matsutomi, Effects of Sharing Farmers' Information using Content Management System, Online Communities and Social Computing, (Eds: A. Ozok, P. Zaphiris), pp.296-303 (2013).

[4] Shiro Takatsudo, Hiromi Murakami and Takeshi Oba, Application of M2M service platform for agriculture ICT, Technical Report of NEC, Vol.64, No.4, pp.31-34, 2011, In Japanese.

[5] Osamu Saito, Fundamental Problems and Strategy about Food-agriculture Alliance-based Food System, Agriculture & Horticulture, Vol.84, No.8, pp.787-793, 2009, In Japanese.

[6] Kiyosada Hayashi, Trade-off in Agricultural Production System, Society for Soil and Water, Vol.27, pp.44-50, 2010, In Japanese.

[7] Teruaki Nanseki, Food Safety and Information Technology, Society of Agricultural Informatics, Vol.17, No.4, pp.161-170, 2008, In Japanese.

[8] Ayumi Shimizu, Tomohiro Uchiyama, The Viability of Farmers Markets' Development by Partnering with Local Restaurants, Agri Economics of Food, Vol.58, No.2, pp.12-19, 2012, In Japanese.

[9] Yuichi Satake, Tomihiro Yamazaki, Using food and agriculture cloud to improve value of food chain, FUJITSU, Vol.62, No.3, pp.262–268, 2011, In Japanese.

Engaging Reflective Thinking in Learning Mathematics Word Problems by Using a Personal Tutoring Agent

Tama Duangnamol†* Boontawee Suntisrivaraporn†
Thepchai Supnithi‡ Mitsuru Ikeda*

†School of Information, Computer, and Communication Technology
Sirindhorn International Institute of Technology, Thammasat University
‡National Electronics and Computer Technology Center
*School of Knowledge Science, Japan Advanced Institute of Science and Technology
d.tama@jaist.ac.jp, sun@siit.tu.ac.th, thepchai.supnithi@nectec.or.th, ikeda@jaist.ac.jp

Abstract

This paper presents a concept of an e-learning system which supports reflective thinking. A Chatbot personal tutoring agent (PTA-Bot) is proposed. Chatbot technology and ontology technology are combined to boost an effective adaptive e-learning system to help students to overcome difficulties in mathematics word problem symbolization and to motivate them to reflect their own thinking during solving problems.

Keywords: Critical Thinking, Reflective Thinking, Chatbot, Adaptive E-Learning System, Word Problem

1 General Instructions

Reflective thinking (RT) is sometime used as synonym of critical thinking (CT), on the other hand, it is a part of the CT process referring specifically to the processes of analyzing and making judgments about what has happened [1, 2]. RT involves personal consideration of one's own learning. It considers personal achievements and failures and asks what worked, what did not, and what needed improvement [3, 4]. It reminds students to think and concern about their own thinking.

Today's world is becoming more complicated. Merely having knowledge or information is not enough. To be effective in the workplace (and in their personal lives), students must be able to solve problems to make effective decisions; they must be able to think critically [5]. To say that, they would have to concern or rethink more about their own thinking before making any decision. As Danielsen mentioned in [6], RT helps students develop higher-order thinking skills by prompting students to relate new knowledge to prior understanding, think in both abstract and conceptual terms, apply specific strategies in novel tasks, and understand their own thinking and learning strategies.

In this research, a case study we mainly focus on is to learn how to symbolise mathematic word problems which is the most difficult implicit step (both for learning and teaching) in solving word problems. These are some reasons of its difficulties in a class room: (i) there is no fixed steps to follows, (ii) its solution is not unique, (iii) students have various capabilities, teacher cannot take care each student personally, (iv) students hesitate to ask a teacher in the class, and the one reason that I absolutely agree with is from [7], (v) students are presented with inconsistent models of problem solving that contradict the logical processes they have learned in other courses and in everyday life.

Students cannot effectively learn how to symbolise word problems by remembering solving strategy step by step, but they have to learn to think critically for reflecting self-thinking to better make a rational decision, even in their dairy life. To engage students to reflect their thinking during practicing on e-learning system, in this research we propose a Chatbot Personal Tutoring Agent (PTA-Bot). PTA-Bot is developed based on Chatbot system by enhancing its semantic power in asking higher-order question and analyzing students' responds for feeding appropriated exercise which is personally suitable for individual student.

In the next section, we present some related works and technologies which are necessary for developing our research. In the Section 3, the architecture design of the system and learning

environment is shown and then activities and exercises which are used in the system are described in Section 4. An example of expected result is presented in Section 5 and in Section 6, the evaluation method is provided. Finally, discussions and future work are expressed in Section 7.

2 Related Works

Ontology is a branch of philosophy which has been widely used in recent years in the field of artificial intelligence and computer and information science, especially in domains such as intelligent information integration, knowledge representation, information retrieval and extraction, work flow and database management systems [8]. There are a number of definitions for describing what is ontology. One of the most widely use is: Ontology is a formal, explicit specification of a conceptualisation [9]. The role of ontology in the field of web-based learning is often underestimated [10]. It has usually been used to describe learners and their profile also to design the learning content and the relation between learning concepts.

In the past ten years, an increasing amount of research has been seen on designing personalised learning systems to deliver learning contents according to the need of particular learner [11]. A personalised e-learning system, namely a personalised intelligent mobile learning system (PIMS), was proposed by Chen et al to promote individual learner's ability of reading English news [12, 13]. It is based on the item response theory that considers vocabulary ability of the learner and learning memory cycle to provide personalised learning [14]. Baylari et al. [15] proposed a personalised multi-agent e-learning system that presents adaptive tests, and acts as a human instructor, and gives the learners a friendly and personalised teaching environment.

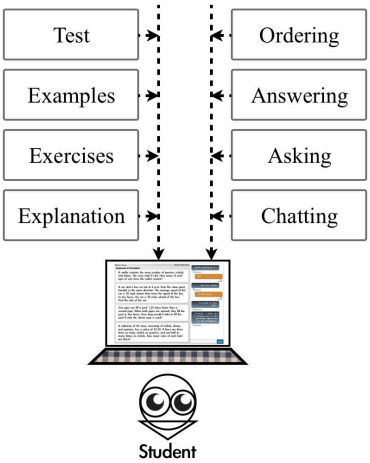
Recently, ontology technologies have become a trend to promote adaptive learning services due to recent advancement of semantic web technologies and tools that allow developers to create meaningful personal user profiles that can assist the learners during the learning process. Vassileva et al. [16] designed learning ontology graphs organizing multimedia learning objects (LOs), and metadata for both the ontology and LOs. Their platforms which consist of an authoring tool, an instructor tool and an adaptation engine controlling the content delivery that adaptable to individual learners. In some works, an ontology was used to model user profiles in different applications like semantic web searching, information retrieval, or natural language processing. Gemmis et al. [17] proposed an extension of the vector space retrieval model in which user profiles are learned by a content-based recommender system. Pan et al. [18] designed a learner ontology and a learning resource ontology for semantic analysis and algorithm for ontology semantic to develop a semantic-based search method for personalised e-learning.

Chatbots are computer programs interacting with users via natural languages. Our choice of choosing a Chatbot is to build a friendly dialog interface, as Zadrozny et al. mentioned in [19] that the best way to facilitate human computer interaction (HCI) is by allowing users "to express their interest, wishes, or queries directly and naturally, by speaking, typing, and pointing". The original Chatbot, ELIZA, was created in the 60's by Joseph Weizenbaum, [20], to emulate a psychotherapist in a clinical treatment. However, Chatbots are not only built to mimic human conversation, and entertain users. It can be useful in several applications, such as, education, information retrival, business, and e-commerce [21]. In last years Chatbot have played a prominent role as human computer interfaces [22]. Knill et al. [23] developed the Sofia Chatbot to assist in teaching Mathematics. The Sofia Chatbot has the ability to chat with users, and at the same time, to chat with other mathematical agents such as Pari and Mathmatica to help in solving Algebra problems. Sofia mainly focuses on mathematics and other common knowledge to make it friendly to use. [23] Teachers could use a Chatbot to look for problems (where students have problems, what questions students ask, and the generated logs file could be accessed to gauge student learning, and students weaknesses) while students use it to solve problems. [21]

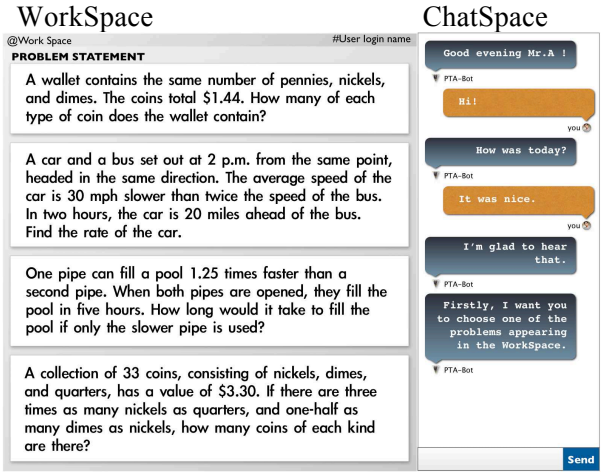
This paper presents a concept idea of designing an alternative adaptive e-learning system, called PTA-Bot, to promote reflective thinking. As mentioned previously, we decided to operate the system via a Chatbot because we would like to simulate friendly learning atmosphere to attract

students' attention. It will be used as a medium to facilitate appropriate learning materials from analyzing conversation dialogs and students' actions or responses.

3 PTA-Bot



(a)



(b)

Figure 1. Learning Environment

Conlan, [24], stated that "e-learning courses typically differ strongly in ethos, learning goals and pedagogical approach whilst learners, even within the same course, may have different personal learning goals, motivations, prior knowledge and learning style preferences." To say that, every individual learns according to their own learning style, needs and interests. Nowadays, personalised learning is an important research issue in the web-based learning field. [11] Accordingly, we aim to create personal tutoring system which supports reflective thinking process to provide appropriate personalised learning-flow to each student individually. In this section, a learning environment of the system and its architectural design are presented.

3.1 Learning Environment

In this learning environment, a student chats with the PTA-Bot on the ChatSpace, and does any activities on WorkSpace in which chatting conversation is related to activities on the WorkSpace, refer to Figure 1(b). In Figure 1(a), while the student chatting with the agent on ChatSpace, the agent can (i) ask the student to do an exercise (the exercise will appear on the WorkSpace), (ii) show some examples to answer to the student question (detailed explanation and some examples will appear on the WorkSpace), (iii) order the student to work on a task provided on the WorkSpace, and (iv) give feedback to the student for his action both in the ChatSpace and WorkSpace.

3.2 Architectural Design

In this research, we propose the Chatbot personal tutoring system which support critical thinking. As shown in Figure 2, the proposed architecture has four layers, compising of learner layer, agent layer, semantic layer, and database layer.

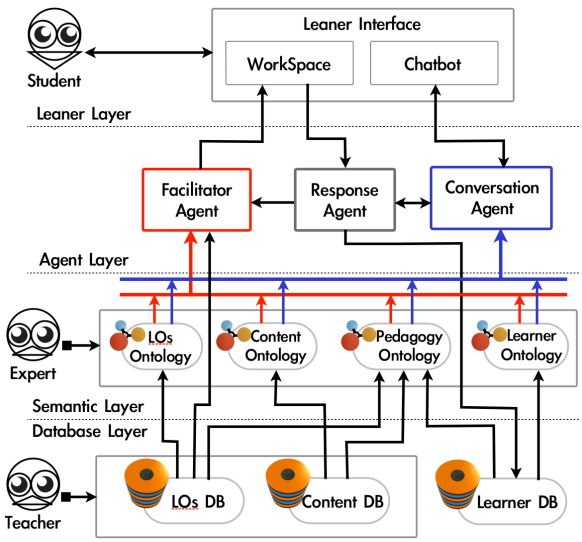


Figure 2. Architecture Design of PTA-Bot.

Learner layer is the layer that interacts with a student, i.e., it is the layer of user interface

which is divided into two main spaces, the Workspace and the ChatSpace, see Figure 1(b).

Agent layer contains three agents: a facilitator agent, a response agent, and a conversation agent.

- *Conversation agent*: This agent drives the Chatbot. It acts as a human tutor that communicates with a student. It is trained (supervised learning) to ask questions to inquire the student ability, to activate student's learning, and to encourage the process of reflective thinking. It answers to the student's questions and give explanation as needed, similarly as existing Chatbots using the knowledge represented in the underlying ontologies and analyzed information from the response agent, shown in Figure 2. It asks the student to work on WorkSpace and gives feedback related to the student action on the WorkSpace. Otherwise, It keeps and stores conversation log and feed to the response agent.
- *Response agent*: This agent records conversation, learning action (such as mouse action), learning duration for each task, learning responses in conversation log, etc. And stores them in the learner database. Moreover, this agent also analyzes the results from student responses, and diagnoses student's learning problems, like a human tutor, and then send information to the facilitator agent and conversation agent to recommend the appropriate learning materials or response messages and interactions to the student.
- *Facilitator agent*: This agent is based on the requests of the response agent which analyzes conversation feed, from the conversation agent, and metadata of LOs and subject content, from LOs and content ontologies, to select the suitable exercises or activities from LOs database.

Semantic layer consists of four ontologies which are provided by domain experts. These ontologies keep semantic structure and relations of concepts in each data knowledge domain;

- Learner ontology: learner's characteristics such as ability and learning styles.
- Pedagogy ontology: pedagogical approaches and instructional designs.
- Content ontology: concepts of subject content, related information, and problem structures.
- LOs ontology: metadata about each LO are preserved.

Database layer contains three databases, a learner profile database which stores user profile, an LOs (learning object) database which store learning materials as learning objects, and a content database which stores the subject content and related information. Data in the learner database is collected from the response agent and being kept under student's login user and the other two databases are provided by teachers.

4 Flow of Learning Activities

In this section, a flow of learning activities is presented. The original problem solving step was proposed by George Polya. It composes of (i) understand a problem, (ii) devise plan, (iii) carrying out the plan, and (iv) examine the solution. The most important step to successfully solve a word problem is to understand the problem, as Polya stated in his book, [25] :

"It is foolish that you answer a question you do not understand. It is sad to work for an end that you not desire".

To understand a problem, a student has to be able to answer these following questions.

- Do you understand all the words used in the stating problem?
- What are you asked to find or show?
- Can you restate the problem in your own words?
- Can you think of a picture or a diagram that might help you understand the problem?
- Is there enough information to enable you to find the solution?
- Do you need to ask a question to get an answer?

You can see obviously, the listed questions can help students to understand a problem by engaging them to remind their thinking process. It

goes with our aim to encourage their own thinking during solving problem. We decide on next learning activities mainly based on the previous questions and referred to problem solving methods mentioned in [25-28].

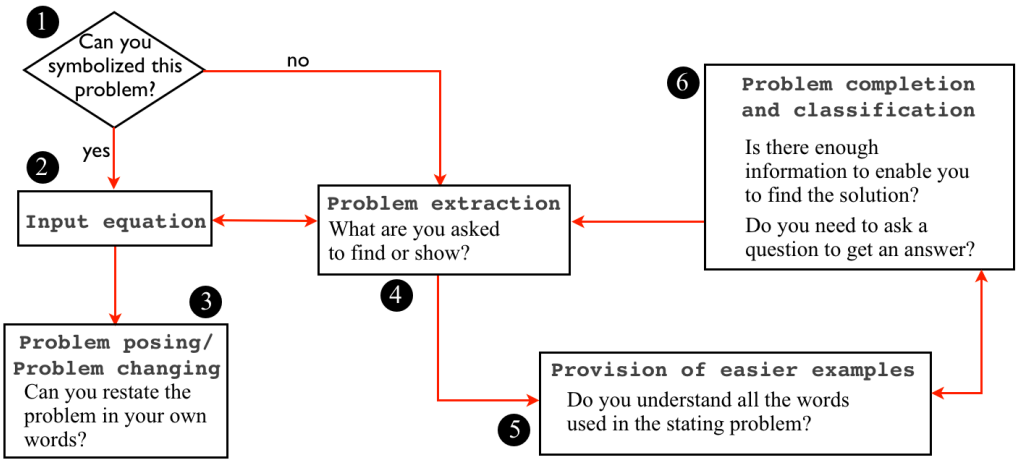


Figure 3. Learning Activity Diagram

***Problem posing and changing*:** Problem posing and changing exercises aim to engage students to rethink and be aware of semantic structure of the problems. A student who can symbolise the problem will be asked to restate the problem from the result equation and be asked to changing some constrains in the problem go meet the expected solution.

***Problem extraction*:** This activity mainly provokes students to be aware of these questions, What is the unknown? What are the data? And What is the condition? In this section a student will be asked to underline important information in the problem.

***Easier example exercise*:** Certain students have no idea about solving word problem they do not know how to start they do not understand the problem's context. In this situation, it is necessary to explain, to give term definition, and to facilitate them some easier or simple examples to build up their basis and to gain more understanding.

***Problem completion and classification*:** In these exercises, a student might be asked to complete incomplete problems and/or give some reasons to classify given problems. These exercises aim to promote students to be aware of problem structures and make them realize about correctness and completeness of a problem, not to just believe that all given problem is correct.

After the student logged-in and accessed to the system, other than chatting with PTA-Bot, the system provides exercises and activities for students. Since, we are concerned about varieties of student's backgrounds and their performances, the system would be able to facilitate adaptively suitable exercises, activities, or contents for each student. From Figure 3, at first, students will be asked to symbolise the problem straightforwardly if they know how to symbolise the problem, they will not get bored at a very basic activities or explanation but they would be offered with more challenging activities to reflect their thinking. Here we use problem posing and problem changing exercise as a reflective thinking tool. If the student could not symbolise the problem correctly, he will be asked to illustrate parts of problems in problem extraction exercises to ensure whether he does really understand the point and structure of the problem. If he can do it really well he will be asked again to symbolise the problem, otherwise, if he is not clear to illustrate parts of the problem, exercises for enhancing his vocabulary and awareness of problem semantic structure, such as exercises on problem completion or problem classification, will be served to the student. Then he would be asked again to distinguish the problem's components before asking to input the solution again to go further to work on problem changing exercises.

5 Expected Results

When a student logins and accesses the system,

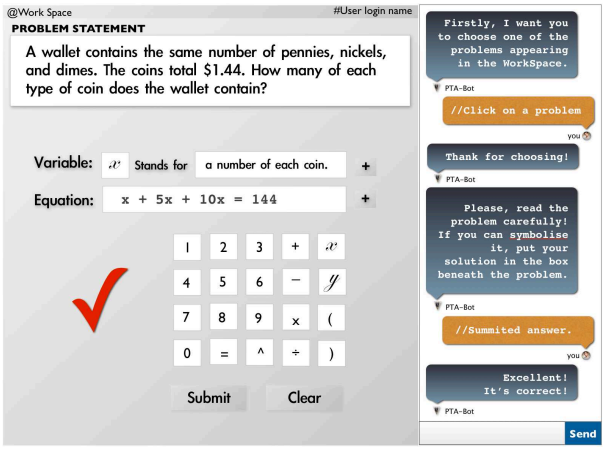
the conversation will be started by the PTA-Bot to make the student feel comfortable to work in the system, as shown in Figure 1(b). As stated in Subsection 3.2, the PTA-Bot communicates with a student on the ChatSpace and delivers conversation-related materials in real-time on the Workspace. From Figure 1(b), in the last chat box from PTA-Bot, it asked the student to select one problem to process, then in the WorkSpace, selection problems will appear. Moreover, when the student performs some action on the WorkSpace, the PTA-Bot is expected to have ability to realize any actions on the WorkSpace by communication among the three agents.
Figure 4 shows the scenario after the student chose a problem. This following dialogue details more about the situation in Figure 4 and what would happen further.

(Continues from Figure 1(b), after the student chose a problem.)
PTA-Bot : Thank for choosing!
PTA-Bot : Please, read the problem carefully! If you can symbolise it, put your solution in the box beneath the problem. (Blank boxes and the input tool appear in the Workspace, see Figure 4(a).)
Student : //input solution
PTA-Bot : Excellent! It's correct!
PTA-Bot : From your solution, which appearing problems. can be symbolised equivalently as that solution? (Show list of questions in the Workspace, see Figure 4(b).)

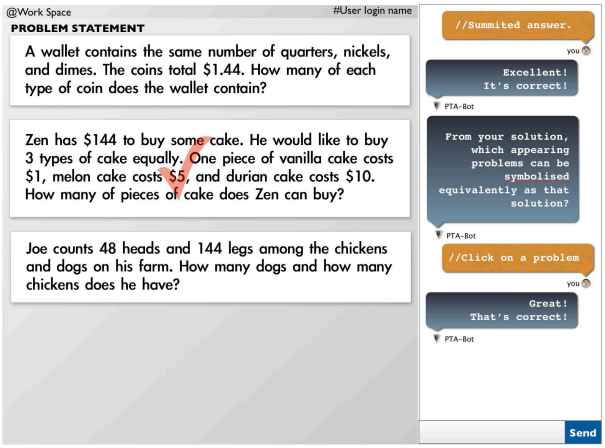
*(If the student chose the correct choice.)
Student : //Click on the correct choice.
PTA-Bot : Great! It's correct!
PTA-Bot : Please use the information appearing in the Workspace to form the problem that contain the same equation. (Information appears in the Workspace.)

**(If the student chose the correct choice.)
Student : //Click on an incorrect choice.
PTA-Bot : Could you please symbolise the problem you just chose? (Blank boxes and the input tool appear in the Workspace.)

From the previous example dialogue, the student is asked to select a question which can be symbolised equivalently as the initial problem to encourage them to rethink about the solution he submitted. Moreover, the student will be engaged to rethink about how to get the solution by Problem posing/changing activities. Such as in (*), in the dialogue, the student was asked to form a new problem which contains the same equation of the main problem. Or in (**), if the student selected a problem which cannot be represented by the same equation, the agent will not be announce his incorrectness but it will ask him to rethink about the solution of the problem he chose.



(a)



(b)

Figure 4. System simulation.

6 Evaluation

The main objective of the system is to promote RT via e-learning system. The consequence of reflective though may motivate students to think more critically that would support them to deeply understand the content they are learning. One way of our choices to evaluate our system performance is to evaluate the

change of students' proficiency, who practice on the system, in solving math word problem using pre-test and post-test comparing with a control group of students, who do not use the system. In pre- and post-testing, other than score comparison, student working-out details will be analysed to observe some critical thinking depositions, such as reasoning, using and mentioning credible source, and relevant thinking to the main point alternative thinking. Students' interaction and conversation logs which are recorded by the system can be also used to analyse some aspects of ongoing interacting situation. This information can be used to observe how students interact with the agent, e.g. do they interact more critically? and does the system help or motivate them reflectively think?

7 Conclusion

In this paper we propose an idea to promote CT in self-learning system. The specific part of CT which is promoted in this work is reflective thinking. A case study that we use in this research is to teach student in solving mathematics context problems using algebraic method in which we mainly focus on the first step of problem solving that is to symbolise a problem, which is the most important and the most difficult step in solving any context problems. To be able to symbolise a problem, it is very important to understand the problem clearly. To understand a problem the questions in Section 4 have been concerned.

As you see, those questions do not only help a student to understand a question, but also engage him to rethink of his understanding. Exercises and activities which are used to support this idea are (i) problem posing and changing, (ii) problem extraction, (iii) giving easier example, (iv) problem completion and classification, and (v) asking higher order questions. To implement our idea, we propose a personal tutoring agent, namely PTA-Bot. User interface of the PTA-Bot has two sections, the WorkSpace and the ChatSpace. The ChatSpace is developed by Chatbot techtology. This is used to create friendly learning atmosphere. The WorkSpace is provided to serve learning materials which are related to the conversation. This mimics the nature of teaching by human that teacher communicate with students and ask the students to do some things. Ontology technology is required to boosted the semantic and consistency performance of the system.

Acknowledgment

This research is partially supported by the SIIT-NECTEC-JAIST dual degree program and the development and promotion of science and technology talents project (DPST scholarship).

References

[1] Dewey, J. (1998). *Reflective Thinking, John Dewey, and PBL*, The Problem Log: Newsletter of PBLNet, vol.3(2), p.3.

[2] Halpern, D. F. (1996). Thought and knowledge: an introduction to critical thinking (3rd ed.). Mahwah, NJ: L. Erlbaum Associates.

[3] Given, B. K. (2002). *Teaching to the brain's natural learning systems*. Washington, DC: Association of Supervision and Curriculum Development.

[4] Vezzuto, L.A. (2005). Reflective Thinking. Orange County Department of Education, retrieved from http://charactered.ocde.us/ICE/lessons_html/reflective.html

[5] Snyder, L. G., & Snyder, M. (2008). Teaching Critical Thinking and Problem Solving Skills. *The Delta Pi Epsilon Journal*, pp. 90-99.

[6] Danielsen, R. (2008). *How Critical is Critical Thinking?*. Clinician Reviews, 18(9), pp. 3-4.

[7] Howard, C. M. (1994). 21st century problem solving, retrieved from http://www.hawaii.edu/suremath/why1.html

[8] Snae, C. and Brueckner, M. (2007). Ontology-driven e-learning system based on roles and activities for Thai learning environment, *Interdisciplinary Journal of E-Learning and Learning Objects*, vol. 3, pp. 1-17.

[9] Gruber, T. R. (1993). *A translation approach to portable ontology specification*. Knowledge Acquisition, vol. 5, pp. 199-220.

[10] Smrž, P. (2004). *Integrating ontologies*

into learning management systems – A case of Czech, presented at the Proceedings of the Workshop on Ontologies, Semantics, and E-learning (WOSE 2004), Berlin.

[11] Yarandi, M., Tawil, A., and Jahankhani, H., (2011). *Adaptive E-Learning System Using Ontology*, Database and Expert Systems Applications (DEXA), 2011 22nd International Workshop, pp.511 – 516.

[12] Chen, C. M., (2008). Intelligent web-based learning system with personalized learning path guidance, *Computers & Education*, vol. 51, pp. 787-814.

[13] Chen, C. M. and Hsu, S. H., (2008). *Personalized iIntelligent mobile learning system for supporting effective english learning*, Educational Technology & Society, vol. 11, pp. 153-180.

[14] Chen, C. M. and Chung, C. J., (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle, Computers & Education vol. 51, pp. 624-647.

[15] Baylari, A. and Montazer, G. A., (2009). Design a personalized e-learning system based on item response theory and artificial neural network approach, Expert Systems with Applications, vol. 36, pp. 8013-8021.

[16] Vassileva, D. and Bontchev, B., (2009). Activities workflow of adaptive courseware design and delivery, in *IADIS International Conference MCCSIS 2009*, Algarve, Portugal, 2009, pp. 223 – 230.

[17] Gemmis, M. d., et al., (2008). A retrieval model for personalized searching relying on content-based user profiles, presented at the Association for the Advancement of Artificial Intelligence.

[18] Pan, J., et al., (2007). A personalized semantic search method for intelligent e-learning, presented at *the International Conference on Intelligent Pervasive Computing*, Washington, DC, USA.

[19] Zadrozny, W., Budzikowska, M., Chai, J., and Kambhatla, N. (2000). Natural language dialogue for personalized interaction. Communications of the ACM, 43(8):116–120.

[20] Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. Communications of the ACM, 10(8):36–45.

[21] Atwell, A. S., (2006). *Chatbots: are they really useful?*, Diabetes Atlas, third edition, International Diabetes Federation.

[22] Augello, A., Pilato, G., Machi, A., and Gaglio, S., (2012). An Approach to Enhance Chatbot Semantic Power and Maintainability: Experiences within the FRASI Project, *Semantic Computing (ICSC), 2012 IEEE Sixth International Conference*, pp.186 – 193.

[23] Knill, O., Carlsson, J., Chi, A., and Lezama, M. (2004). An artificial intelligence experiment in college math education. Preprint available at http://www.math.harvard.edu/_knill/preprints/sofia.pdf.

[24] Conlan, O., (2006). *The multi-model, metadata driven approach to personalised elearning services*, Doctor of Philosophy, Department of Computer Science, University of Dublin,Trinity College.

[25] Polya, G., (1957). *How to Solve It : A New Aspect of Mathematical Method*, 2nd ed., Princeton University Press.

[26] Yamamoto, S., Waki, H., and Hirashima, T., (2010). *An Interactive Environment for Learning by Problem-Changing*, ICCE2010, Nominated for Best Technology Design Paper Award(2010.12)

[27] Charles, R. I., (2012). *Solving Word Problems : Developing Students' Quantitative Reasoning Abilities*, Research Into Practice MATHEMATICS, Pearson Education, Inc.

[28] Karnjanamayoon, S., (2002). Solving word problem, *Journal of Science Mathematics and Technology*, IPST, vol30(116). (in thai) (Author)

A Virtual Environment for English as Foreign Language Learning Platform (veEFL): Applied "Single Idea of Concept" to Improve Writing skill of Low English Proficiency Students

Nattapol KRITSUTHIKUL†,*,** **Shinobu HASEGAWA‡,***
Cholwich NATTEE† **Thepchai SUPNITHI****

†Information, Computer, and Communication Technology (ICT), SIIT, TU, Thailand
‡Center for Graduate Education Initiative (CGEI), JAIST, Japan
*School of Information Science (IS), JAIST, Japan
**National Electronic and Computer Technology Center (NECTEC), Thailand
nattapol.kritsuthikul@gmail.com

Abstract

In order to assist EFL students with low English proficiency in learning writing skill, we propose a framework of a virtual environment to evaluate common errors that the students often conduct in writing essay. The system is a service applied in platform linked with other NLP web services to help with language analysis. The system mainly focuses on finding the writing errors related to semantic meaning selection, incorrect structure to indicate the intended meaning, non-smoothing sentence in topic, and etc. The system improves the students' writing skill by providing questions relating to the matters they are writing.

Keywords: Single Idea of Concept, Unified Paragraph, English as Foreign Language Learning (EFL), Writing Skill, Low English Proficiency, Natural Language Processing (NLP), Web Service

1 Introduction

In language learning, four skills - reading, writing, speaking and listening, are required to be learnt and practiced. Among those skills, writing is skill for students to be acquired because it is mandatory to express students' thought in documents. Writing skill is one of the most troublesome skills in teaching [1] since it is difficult to measure on how to correct the written sentence matched with students' thought. Apart from grammatical error, meaning and conceptual are more difficult to be corrected. For non-English-native students, the usage of words may be ambiguous due to the low efficiency and lack of understanding in word senses.

1.1 Lower-secondary school education in Thailand

In Thailand, English studying program in lower-secondary school education (Grade 7-9) isn't arranged as a study program in major likes in high school education (Grade 10-12) or University. All Students in lower-secondary school will study all the same curriculum in their own level. On the other hand, Students in high school education have to select what major they would like to study such as "Science and Mathematics program" or "Liberal Art in English program".

Students who study English in lower-secondary school education will be taught English as one subject including all basic skills in one unlike students studying English in high school education will focus on specific skills in English such as English Writing, and English Reading. Teachers teaching English in lower-secondary school will have a mini test which is divided a test to examine each English skill in their students like Grammar, Reading, and Writing.

1.2 Low English Proficiency Students

Low English Proficiency Students (LEPS) will be the students who have under average score in English writing skill which the English teacher may choose those LEPS. In versus, the students who have good or excellent score in English writing skill are High English Proficiency Students.

LEPS are not able to develop their English skills independently. When they make some mistakes on their writing, they cannot recognize where the mistake are and why they make them wrong. They need an English teacher to suggest and explain about the reason of making mistakes. These will help them to develop their English skills gradually. The difference between High English Proficiency Students and the LEPS are the ability to study. The LEPS, they must ask their teacher more often than the High English Proficiency Students do in order to develop their learning skill step by step; whereas, the High English Proficiency Students can learn their mistakes without other helping. Thus, veEFL is built to assist improving English writing skill for the LEPS.

The best way of learning is to enable the student to develop his thinking by himself. Develop thinking is very important element to apply it into writing ability. When the student would like to share his idea, it is very essential to know how to begin, continue and end up the story. Some of the student s have some difficulties on how to continue the story. As provided a writing assignment about "What is your favorite hobby?", for example, most of the learners…". On the other hand, the LEPS may come to impasse to construct sentence (or expand sentences) smoothly.

However, the problem is how they continue the rest of their favorite hobby. This is very necessary to help the student to develop his thinking and to write more sentences in his writing. Therefore, veEFL provides the students with questions relating to the matters they are writing. veEFL will generate the questions from the student's input sentence at real time. veEFL will generate the questions from the student's input sentence at real time. veEFL will analyzing structure of the input sentence and extracting into keywords. By this method, the student will expand the idea from the generated questions from veEFL. The result is that they can write more sentences related to the topic and understand by themself how to construct more sentences relating to the keyword.

To implement the above-mentioned tools, there are several necessary Natural Language Processing (NLP) resources and tools such as WordNet [2] (for synonymous and hierarchical meaning), FrameNet [3] (for theme accordance), Syntactic Parser (for analyzing main verb, subject and complement), and so on. To facilitate the writing assistant tools, Web Service technology is introduced to link existing shared veEFL Web Service Resources.

2 Common Errors in EFL Writing skill of Low English Proficiency Students

Thai Grade 7 students have started studying English from vocabulary, part of speech, type of sentence and tenses respectively. For Thai Grade 8-9 students, they have practiced English writing. If the students want to create good English Writing skill, they need to know how to apply Unified Paragraph idea and Paragraph Composition together. This skill will make the students create their own ideal writing essay.

There are three obstacles of writing Unified Paragraph as follow: (1) Semantic Meaning Selection, (2) Incorrect structure to indicate the intended meaning, and (3) Non-smoothing sentence in topic. The last problem (Non-smoothing sentence in topic) is also a problem of creating Paragraph Composition.

2.1 Semantic Meaning Selection

Semantics is the study of the meaning of linguistic expressions. It concerns about the relation among words, phrases, signs and symbols. So, the problem on Semantic Meaning Selection is the problem of choosing an inappropriate word applied in a context in order to create a misunderstood translation and leads difficultly to find the intended meaning in the sentences from the written content.

Most of Low English Proficiency of Thai students often uses Thai to English translation as a pattern applied in English writing. This leads them to use a word to word translation. As a matter of the fact that the simpler the word is, the more senses the word can represent, e.g. Thai verb "เขียน (write : verb, draw : verb, paint : verb)" can be used collectively with nouns such as "หนังสือ (book : noun)", "เรื่อง (story : noun)", "รูป (picture : noun)", etc. while the translation must vary to the object. Thai students often incorrectly use the phrase "write a picture" instead of "draw a picture" from lacking of understanding of word sense.

In such circumstances, the teachers may find it particularly difficult to guess the accurate

meaning of students' thought without having personal discussion. In order to expedite the study of writing, tools for scoring are essential for writing a given topic and suggesting appropriate writing concept

To resolve this problem, Thai knowledge should be applied as follow are needed (1) the most frequently misused word to word translation from Thai to English, (2) Thai WordNet [4], and (3) Thai FrameNet [5].

2.2 Incorrect structure to indicate the intended meaning

The structure of a sentence is a group of words gathering as a sentence in order to create the intended meaning. The structure is related to grammars. Incorrect structure is also grammar mistakes. Incorrect structure is more likely to express the wrong intended meaning.

In Thai language, there is no rule of verb inflection. They can use all the same verb in any time of situation but in English, we have to change the verb form to suit with the time of situation happens. Though this type of error is not serious mistakes, it can create the problem of misunderstanding.

This is an example sentence of “no rule of verb inflection” that reflects to this problem, “Yesterday, I go to school”. From this sentence, we all know that the word “Yesterday” is used to show that the situation was taken in the past but the verb in this sentence “go” is in present form. Therefore, the verb “go” must be changed to be “went”, past form of the verb “go”. The correct sentence must be "Yesterday, I went to school" in order to tell that the situation was already happened and finished.

This type of error is a common problem for Thai students because most of them usually think in Thai words and then translate into English. For example “เมื่อวานฉันไปโรงเรียน”. Due to lack of clear marker for word and sentence boundaries in Thai. We need to applied Thai word segmentation. The result are “เมื่อวาน (yesterday)”, “ฉัน (I)”, “ไป (go)”, and “โรงเรียน (school)”. From translating Thai to English as basis, we can see the problem of wrong structure which can create wrong meaning. “Yesterday, I go school” - Translation from Thai to English. The verb tense in this sentence is wrong and there is a missing of the preposition. The verb tense should be changed to be in past form and the preposition should be added also. The correct sentence should be “Yesterday, I went to school”. However, Semantic Meaning Selection problem is difference because it doesn't make the grammatical errors in the sentences

2.3 Non-smoothing sentence in topic

When student write an essay, there are many groups of sentences gathering together as a paragraph in order to describe what the content is. Every single sentence in a topic will be arranged to generate the student's idea to the teacher's understanding. Smoothing sentence can help the reader easily sequence the concept (Topic). Therefore, we can imply that non-smoothing sentence in topic can lead both the student and the teacher misunderstood the intended content.

For example of the sequence of time. There are three sentences gathering together to tell the story:

1. I go to school
2. I study Math in the afternoon
3. I play football in the morning

From these three sentences, we can assume that “I go to school” must be the first sentence of this story because you have to go to school first in order to study and do some activities in the school. Then, the sentence number 3, “I play football in the morning”, will be the second sentence because we can recognize the word “in the morning”, the sequence of time happened before the word “in the afternoon”. Thus, the sentence number 2 will be the last one.

For example of non-coherence. There are three sentences gathering together to tell the story: (1) “I go to school.” (2) “My father likes fishing.” (3) “My friend likes playing football.” from these sentences, none of these sentences is related together.

Lastly, writing smoothing sentences can help not only the students organize and generate their idea but also the teachers understand the story easier. To resolve this problem, Single Idea of Concept is applied to assist the student to writing more smoothing.

3 Single Idea of Concept

Unified Paragraph [6] is a paragraph which focuses on explaining one idea in one paragraph so Unified Paragraph is a part of “Single Idea of

Concept". Unified Paragraph should consist of three parts; introduction, body and conclusion. All these three parts must be written with only one idea. A paragraph should consist of more than 200 words up which may be complicated for the student to write many sentences to be a paragraph. Due to the ability of the student, not many of them can write up to 200 words in English. Therefore, we would like to apply only the concept of Single Idea to this method. The aim of this method is to help the student developing his mind and focusing on what he is writing so Single Idea of Concept is more appropriate to apply in it.

Single Idea of Concept is a basement of writing ability which helps the student not to miss his objective of writing so Single Idea of Concept is a concept of thinking in one story at a time. Single Idea of Concept is to enable applying with LEPS because it doesn't expect the student to write it as a full paragraph. Most of LEPS are not able to write a full form of paragraph which is required more skillful in explaining English with long sentences or many short sentences to be a paragraph. Therefore, we would like to use Single Idea of Concept to help the student to express his mind by not focusing on the correct form of a paragraph in order to avoid narrowing the student's idea. Single Idea of Concept will help the student developing his mind on what the content of his story should be discussed on. Single Idea of Concept will help the student by focusing on one keyword of the story and then extending more questions related to the keyword.

For example, the student has an assignment about telling about your hobby. Single Idea of Concept will be used to help the student of what his hobby is. The student will be given the question relating to the word "hobby" such as "What is your hobby?", "Where do you do your hobby?", "When do you do your hobby?", "Who do you do your hobby with?", and "Why do you like your hobby?"

Then, the student can begin his story by answering those above questions.

For example,

My hobby is playing football. I play football at garden yard. I play football after finishing school. I play it with my father. I like play football because it is fun and good for health.

Though, this is not a full form of a paragraph but it helps the student to express his idea and write more sentences about his story. Single Idea of Concept is to use to help the student developing his mind with one idea in one story. It is a basic of good writing to write the story straight to the point of what he would like to talk about.

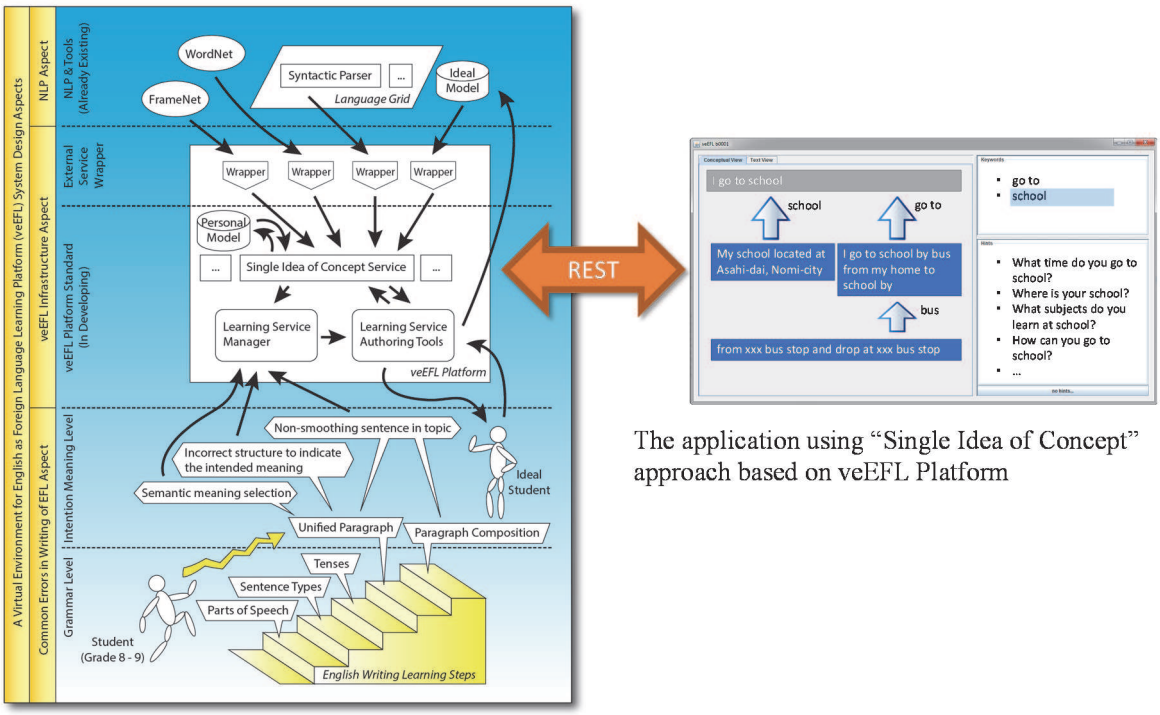
Coherence and cohesion is one of technical elements to create a good comprehensive paragraph. Coherence [7] is to use semi-technically of the way in which the content of connected speech or text hangs together to make the content meaningful. Cohesion [7] is to connect a text together linguistically. Therefore, coherence is to connect the content of text smoothly and create good comprehension. Cohesion is to connect the text linguistically whether it creates smooth good understanding or not.

Single Idea of Concept is different from coherence and cohesion because of the aim to apply on it. Single Idea of Concept is to focus on one single idea in the context by not focusing on the term of grammatical construction and good combination to create smooth comprehension. According to our target group, Single Idea of Concept is to help them get free from the struggle of inability to write. Single Idea of Concept is not used to point to correct grammatical construction, create good sequence and pass on knowledge smoothly. We use Single Idea of Concept to encourage whom are not able to write and continuous their writing to be able to write more sentences. Coherence and cohesion acquire more skillful from the writers to write correct grammatical construction and link those sentences with good explanation to the readers.

In summary, Single Idea of Concept is different from cohesion and coherence because of the aim of it and the suitability to the target group (LEPS). Single Idea of Concept aims to encourage the target group whom has not much knowledge on English to be able to write and continue their writing by not pushing on correct grammatical construction and good sequence to pass on the content smoothly which require more advanced knowledge in English.

4 System Architecture

Developing innovative products needs a lot of time and budget consumption. If there is not good



The application using "Single Idea of Concept" approach based on veEFL Platform

Figure 1. System Architecture and the application based on veEFL Platform

enough platform, it is very difficult to accomplish the innovative service. The current trend on research is to aim collaboration among other parties and to develop a large scale project. The management of project is to integrate all components and become a challenging task.

We propose a new web service platform named "a Virtual Environment for English as Foreign Language Learning Platform" (veEFL) that is compliant with WS-I. WS-I [8] is a specification for maximizing compatibility with Web Service Platforms. This platform will support Document/literal wrapped SOAP message binding style, REST with Token-based Authentication for easy to consumed web service, and HTML5 WebSockets [9] for two-way communication over the Internet. The System Architecture and the application based on veEFL Platform show in Figure 1.

This system is designed by concern of 3 difference aspects as follow (1) Common Errors in Writing of EFL learner aspect (2) Infrastructure aspect, and (3) NLP aspect. veEFL is acted as a service broker to synergize all aspects together.

4.1 Common Errors in Writing of EFL learner aspect

This view point focuses on how these platforms can help Thai students who have Low English Proficiency.

English writing assignment for Thai Grade 8-9 students, the teacher often specifies writing topics for students. Writing topics will be simple open-ended questions about the life style of the students.

Examples of simple assignment for English writing:

- What is your hobby?
- What do you do on your free time?
- What would you like to be in the future?
- What will you do when you finish school?
- How do you go to school?
- Tell me about your school.

Though these questions are simple, they are still difficult for LEPS. They are more likely not capable to write many sentences joining as an essay. They are capable to write it, nonetheless; the sentences they wrote won't be linked up as a

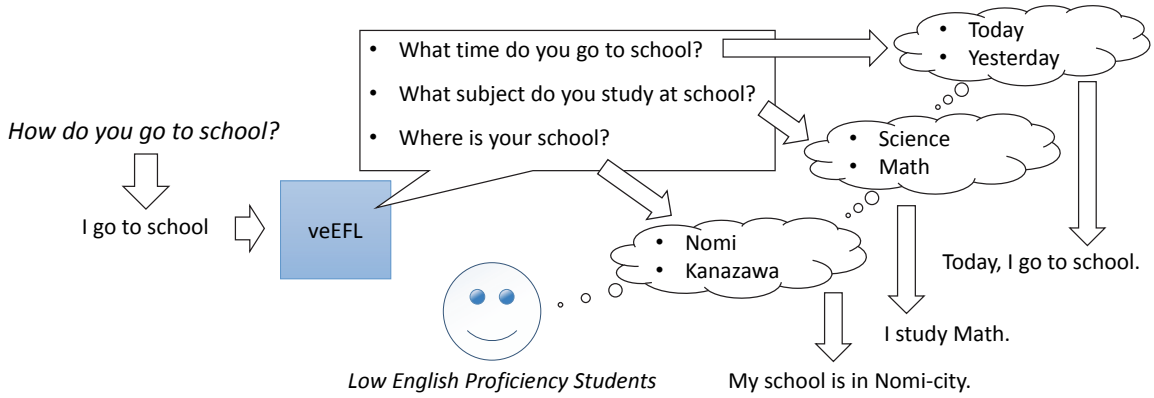


Figure 2. An example of apply a story-line concept as to be a Single Idea of Concept

same story line. They aren't capable to write sentences to clarify details. Most of their writings have a problem of non-smoothing sentences in a topic because of not understanding of Single Idea of Concept.

A guideline to solve student's weakness, non-smoothing sentences in a topic, is to build a tool for motivating LEPS to apply a story line concept as to be a single idea concept. This tool is a writing guideline as a story line. The initial concept is shown in Figure 2.

This system will receive input as a sentence. The system will assume the first sentence as a topic sentence. For an example, "How do you go to school?" The system will analyze the elements of each sentence such as "go to (v)" and "school (n)". Then, the system will suggest the story which is related to the word appeared in a topic sentence. For example, the system will present the contents which are associated with the word "school" in order to be a guideline for the students to write on their story line such as

- What time do you go to school?
- Where is your school situated?
- What subjects do you study at school?

However, in some questions the system can give a possible answer such as "What subjects do you study at school?" so the answer will be the name of subjects. This will help the students to write more details (sentences) on their story. The Example of application based on veEFL Platform show in Figure 1 (in right side).

The left panel is shown the conceptual dependency of initial sentence from student, "I go to school". The top-right panel is shown the extracted keywords from initial sentence. The application can extract into 2 keywords: "go to" and "school". The highlight on "school" keyword is an active keyword. The bottom-right panel is shown generate questions relating to an active keyword. The student will see the questions and expand his idea to write other sentences.

4.2 veEFL infrastructure aspect

veEFL is a combination of several technologies for assisting English learning. It pores on increasing Thai student's writing skill.

The basis of the system is the development of unique task based service template (veEFL Resource Template) and sequencing them to accomplish the final objective. In the beginning phase, the initial service template would be a service for practicing basic skills such as translation service template.

The main system consists of (1) Virtual Environment Learning Platform Service Infrastructure Standard, (2) Learning Service Manager, (3) Virtual Learning Composite Workflow Builder, and (4) Learning Service Authoring Tools.

A Virtual Environment Learning Platform Service Infrastructure Standard is the one for service template on learning. It focuses on fast, simple and reusable generation of service template (veEFL Resource). The service template is designed to be connected as JSON style REST web service and supported to HTML5 WebSockets Technology for variety of service available such as stateful service, broadcast

streaming service, real-time collaboration service and so on.

A Learning Service Manager is a controlling system of service flow and user permissible activity. Actually, veEFL system is coordinated with many veEFL Resource. veEFL Resources are the registration of User Account. The User Account of veEFL Resources will be the owner of his or her own resources. The owner can allow other User Account to reach the owner's resources. Many kinds of veEFL Resources are provided by veEFL platform, e.g. veEFL Atomic Web Service Resource – the resource that can be used as Web Service, veEFL Composite Web Service Resource – the resource that allow sequencing veEFL Web Service Resource, and so on.

A Virtual Learning Composite Workflow Builder is a connecting system to compose two or more existing services into complex service. It focuses on the simplicity of usage and instant composing existing services.

Learning Service Authoring Tools are the ones for composing services as components like Rich internet application (RIA). This tool is used to generate RIA application based on veEFL platform.

Finally, there are some reasons to use following technologies to support veEFL platform

- REpresentational State Transfer (REST)
- Token-based Authentication
- HTML5 WebSockets

4.2.1 REpresentational State Transfer (REST)

SchoolNet is a Thai Government's project supporting internet availability to cover all schools in Thailand. This project is guaranteed the minimum standard of V.92 at 56 kbs over telephone network.

We should be concerned with the internet connection speed of SchoolNet. Thus, REST is more suitable than SOAP because of its lower overhead.

4.2.2 Token-based Authentication

WS-Security [10] is a standard security mechanism for SOAP web service. Unlike SOAP, REST is lack of standard mechanism for security and authentication. The platform needs a proper security mechanism and the best choice is Token-based authentication. It is easy to implement and can be harmonize with REST over HTTP. This technique will help user to consume Web Service easier. In case of higher security level, REST communication over a computer network is needed, HTTPS can be applied.

4.2.3 HTML5 WebSockets

Two ways communication between teacher and student interaction is ordinary task on learning. It means that the learning platform needs to provide efficient real-time, bi-directional, and full-duplex communications to serve user interaction. HTML5 WebSockets (or WebSocket in short) is a new W3C standard as part of HTML5 specification that enables all required functions needed by the platform.

4.3 Natural Language Processing (NLP) Aspect

To efficiently manage the English content, NLP tools are mandatory to handle language information. A Part-of-Speech (POS) tagging tool and a syntactic parser are subjected to syntactic analysis while semantic information is taken care by available English WordNet and FrameNet. These tools will be used to analyze students' writing for correctness, compatible meaning and collocation of the sequence of the sequence of lexical constituents

6 Conclusion

In this paper, we propose a framework of A Virtual Environment for English as Foreign Language Learning Platform infrastructure (veEFL) as a new Web Service infrastructure platform. We apply the existing service of NLP tools with the idea of realizing a meaning of writing essay to find the connection between written content. We plan to develop a scoring system to evaluate the smoothness of topic between written sentences of EFL students who are Low English Proficiency, and develop a system to suggest the appropriate way on writing.

Acknowledgments

Special thanks to Dr. Krit Kosawat for helpful discussions, Dr. Marut Buranarach and Mr. Tanate Ruangrajitpakorn for insightful comments. This research was conducted under a grant in the SIIT-JAIST-NECTEC Dual Doctoral Degree Program.

References

[1] Pawapatcharaudom, Ratana. (2007). *An Investigation of Thai Students' English Language Problems and Their Learning Strategies in The International Program at Mahidol University* [Case study]. King Mongkut's Institute of Technology North Bangkok, Thailand.

[2] Princeton University (1985). *WordNet: An Electronic Lexical Database*. [On-line]. Available: http://wordnet.princeton.edu/

[3] International Computer Science Institute in Berkeley, California (1997). *FrameNet*. [On-line]. Available: https://framenet.icsi.berkeley.edu/

[4] Leenoi, Dhanon. (2008). *The Construction of Thai WordNet of 1st Order Entity Common Base Concepts Using a Bi-Directional Translation Method and With Dictionaries of Different Compilation Approaches* (Master's thesis). Chulalongkorn University, Thailand.

[5] Leenoi, Dhanon. Jumpathong, Sawittree. and Supnithi, Thepchai. (2010). *Building Thai FrameNet through a Combination Approach*. Proceedings of IALP 2010 (pp. 277-280). Harbin, China.

[6] Crossley, S. A., Dempsey, K., & McNamara, D. S. (2011). Classifying paragraph types using linguistic features: Is paragraph positioning important?. *Journal of Writing Research*, 3(2), 119-143.

[7] Matthews, Peter. (1997/2007). *The Concise Oxford Dictionary of Linguistics (2nd edition.)*. Oxford University Press.

[8] The Web Services-Interoperability Organization (WS-I) (November 9, 2010). *WS-I Basic Profile Version 2.0*. [On-line]. Available: http://ws-i.org/Profiles/BasicProfile-2.0-2010-11-09.html

[9] Internet Engineering Task Force (IETF) (2001). *HTML5 WebSockets*. [On-line]. Available: http://www.rfc-editor.org/rfc/rfc6455.txt

[10] Web Service Security (WSS) (February 1, 2006). *Web Services Security: SOAP Message Security 1.1 (WS-Security 2004)*. [On-line]. Available: http://www.oasis-open.org/committees/download.php/16790/wss-v1.1-spec-os-SOAPMessageSecurity.pdf

[11] Kritsuthikul, Nattapol. Nattee, Cholwich. Supnithi, Thepchai. & Hasegawa, Shinobu. (2012). *A Virtual Environment for English as Foreign Language Learning Platform (veEFL): Common Errors in EFL Writing of Low English Proficiency Students Infrastructure*. Proceedings of SIG-ALST 66 (pp. 31-34). Kanagawa, Japan.

[12] Kritsuthikul, Nattapol. Hasegawa, Shinobu. Nattee, Cholwich. & Supnithi, Thepchai. (2013). *A Virtual Environment for English as Foreign Language Learning Platform (veEFL): Applied "Single Idea of Concept" to Improve Writing skill of Low English Proficiency Students*. Proceedings of ICCE 2013. Denpasar Bali, Indonesia.

An Image Creating System based on Kansei of Natural Language Sentences

Miki Ueno† **Akihiko Yasui**‡ **Naoki Mori**† **Keinosuke Matsumoto**†
†Graduate School of Engineering ‡College of Engineering
Osaka Prefecture University
1-1 Gakuencho, Nakaku, Sakai, Osaka 599-8531, Japan
ueno@ss.cs.osakafu-u.ac.jp

Abstract

Recently, lots of researchers in computer science field pay attention to the personal process called "Kansei" that is human's mapping rule between an external stimulus to an emotion. However, it is difficult to represent the "Kansei" process as a computer algorithm. In this paper, we propose the novel method of utilizing "Kansei" in computer applications that produces fractal images based on numerical vector of user's "Kansei". We assign emotional labels to images according to questionnaire in order to map user's "Kansei" to certain images. The computer experiments are carried out to confirm the effectiveness of proposed method.

Keywords: Automatic Drawing Picture by Natural Language, Fractal Figures, Kansei Engineering, Sentiment Analysis

1 Introduction

The automatic picture drawing based on human emotions/thoughts by computer is one of the most important and difficult topic in information science field. Several researchers have focused on the relation between human emotion/thoughts and representation. There are lots of variation of representation, we focus on picture representation in this study.

There is well known interesting result of relation between specific sound and specific object in psychological field called buba/kiki phenomenon[1]. The result of this phenomenon shows that 95% of English-speaking adults matched "bouba" with around, amoeboid shape and "kiki" with an angular figure. Furthermore, the game is designed to assess 4 kinds of known mapping and to check differences between adults and children [2].

From the viewpoint of the way of communication, the method that translates a picture made of pictograms into a sentence and a sentence into symbolic pictures has been proposed[3]. In this method, the most important requirement is that the produced pictures must have only one meaning because target users of this method are people who can not read, but can understand significant information such as disaster prevention through the picture.

In order to introduce human emotions to computer applications, rule-based affect analysis model has been proposed[4; 5]. To simplify the model and make analyzing emotions easy, sentiment analysis has also been proposed. Sentiment analysis is a method which is classifying the polarity of a given text based on the document or sentence into positive or negative by means of utilizing sentiment dictionary[6] . However, there are few research about quantitative analyses to translate human mind into picture image [7]. To solve this problem, we propose a novel method which translates a user input sentence into a picture by evaluating user's "Kansei". We also made standalone application of the proposed method to show the experimental results.

In this paper, we propose the method of producing images by user's inputs in section 2. We show how to assign emotional labels in section 3. Computer experiments are described in section 4, while section 5 shows the method of expanding dictionary. Section 6 introduces application example using the proposed method. Finally, in section 7, we present the conclusions of this study.

2 Proposed Method

2.1 Input

Though humans have several expression to represent their *kansei*, we assumed that user's input is natural language sentence because natural lan-

guage is one of the most common communication method in the internet environment.

2.2 Emotional Output

In this paper, we define *kansei* which varies from person to person as rules of mapping an external stimulus to an emotion. We define *kansei* of input s as numerical vectors. In this paper, we define 2 types of numerical vectors $\boldsymbol{\epsilon_1}$ and $\boldsymbol{\epsilon_2}$ in order to represent *kansei*. Though it requires vast amount of elements to represent human affects, to make problem simple, elements of $\boldsymbol{\epsilon_1}$ represent the only intensity of positive and negative. On the other hand, elements of $\boldsymbol{\epsilon_2}$ represent the intensity of 6 basic emotions, $joy, anger, sadness, hate, fear, surprise$.

The emotional intensity is calculated by search engine result. We select the 6 typical words which have strong relation to 6 emotion categories above. These words are defined as *sentiment words*. Utilizing result of web search based on these *sentiment words*, we obtained $\boldsymbol{\epsilon_2}$. An emotional intensity is calculated by following method:

Emotional Intensity : Utilizing the number of result pages in Yahoo! JAPAN keyword search, the emotional intensity e_i^w is determined by morpheme w and *sentiment words* χ_i in $\boldsymbol{\chi}$ by following equation:

$$e_i^w = [\mathrm{search}(w \mathrm{AND} \chi_i)]/[\mathrm{search}(w)] \quad (1)$$

,where $i = 1 \cdots 6$ represents one of 6 emotions {1:joy, 2:anger, 3:sadness, 4:hate, 5:fear, 6:surprise}. The emotional intensity e_i of s is obtained by $\sum_w e_i^w$.

Search(x) is the number of result pages of query x. "AND" means AND search by several queries.

2.3 Emotional Output Mapping

Proposed algorithm of mapping input s to $\boldsymbol{\epsilon_1}$ and $\boldsymbol{\epsilon_2}$ is as follows.

2.3.1 *Kansei* Vector $\boldsymbol{\epsilon_1}$

1. Obtain user's input s.
2. The number of morphemes N without particles, auxiliary verbs, unknown words and symbols is obtained after applying morphological analysis by Sen. Each morpheme is decided $w_i (i = 1, 2, ..., N)$.
3. Let j = 1. W_p = W_n = 0.
4. Let t = 0. If $j > N$ then go step 7, if $j \leq N$ then set w_{tmp} = w_j.
5. If w_{tmp} is contained in the sentiment polarity dictionary and the sentiment polarity is positive then W_p = W_p + 1. If the sentiment polarity is negative then W_n = W_n + 1. After that, set $j = j + t + 1$ and go step 4. If w_{tmp} is not contained in the dictionary then set $t = t + 1$ and go step 6.
6. If $j + t$ = N + 1 then go step 7. If $t < 3$ then word combined w_{tmp} and w_{j+t} is defined as w_{tmp} and go step 5. If t = 3 then set $j = j + 1$ and go step 4.
7. Finally, *kansei* vector $\boldsymbol{\epsilon_1}$ is set to $(\frac{W_p}{N}, \frac{W_n}{N})$.

2.3.2 *Kansei* Vector $\boldsymbol{\epsilon_2}$

By equation (1), *kansei* vector $\boldsymbol{\epsilon_2}$ is calculated by input s as follows.

1. Obtain user's input s.
2. The number of morphemes N without particles, auxiliary verbs, unknown words and symbols is obtained after applying morphological analysis by Sen.
3. Set *sentiment words* vector $\boldsymbol{\chi}$ = $(\chi_1, \chi_2, \cdots, \chi_6)$ and *kansei* vector $\boldsymbol{\epsilon_2}$ = $(e_1, e_2, ..., e_6)$, where $i = 1 \cdots 6$ represents {1:joy, 2:anger, 3:sadness, 4:hate, 5:fear, 6:surprise}. Initial value of e_j is set to 0. χ_i is decided manually based on emotion i.
4. Let $p = 1$.
5. Search in Yahoo! Japan Keyword by w_p as query and define the number of result pages as W.
6. Define $X_j (j = 1, 2, ...6)$ as each number of page results obtained by AND search with w_p and χ_j.
7. Let q = 1.
8. Let $e_q = e_q + \frac{X_q}{W}$.
9. If $q \leq 6$, let $q = q + 1$ and go step 8.
10. If $p < N$, let $p = p + 1$ and go step 5.
11. Finally, let *kansei* vector $\boldsymbol{\epsilon_2}$ = (e_1, e_2, ..., e_6).

2.4 Representation of Emotional Output

In the proposed method, colors of image are determined by $\boldsymbol{\epsilon_1}$ and the pattern of image is determined by $\boldsymbol{\epsilon_2}$. The pattern of image is created by the Julia set[8] which is one of the popular fractal patterns.

2.4.1 Settings

In this paper, we use an image for an example of emotional output. A certain situation reminds human of some images such as symbolic images, scenery images and abstract paintings. Computers recognize the image as only matrix of color value on each pixels. We set the width and height of images to 500 pixels and color model to $\boldsymbol{RGB}$.

2.4.2 The Julia Set

We use the Julia set to represent emotional output. Following equation of complex function of complex number z is defined first.

$$f(z) = z^2 + C \tag{2}$$

Constant C is defined as follows.

$$C = p + qi \tag{3}$$

Define following recurrence formula:

$$Z_{n+1} = f(Z_n) \tag{4}$$

To obtain Julia set, judge whether the $\lim_{n\to\infty} Z_n$ is convergent or not. The border between convergent and divergent is the Julia set.

2.4.3 Condition of Convergent

In theory, to judge whether a complex number is convergent or not, we check the limiting value of Z is required. However, in the proposed system, if absolute value of Z_n is over 2, we regard the value as divergent.

1. Let n = 0
2. Let $z_0 = 0$
3. Let $n = n + 1, Z_{n+1} = f(Z_n)$.
4. If n $\leq$ 2000 then n_{div} = 0 and go to step 3.
5. If $|Z_n| > 2$ then $n_{div} = n$.

2.4.4 Drawing Algorithm

Each $C_i (i = 1, 2, ..., 6)$ relates to a label of 6 emotions.

1. Obtain $\boldsymbol{\epsilon_1}, \boldsymbol{\epsilon_2}$.
2. We set the initial value of Julia set for 6 different emotions as follows:

$$C_1 = 0.3 \tag{5}$$
$$C_2 = -1 - 0.27i \tag{6}$$
$$C_3 = 0.4 + 0.1i \tag{7}$$
$$C_4 = -0.74 - 0.18i \tag{8}$$
$$C_5 = -0.06 - 0.68i \tag{9}$$
$$C_6 = -1.37 \tag{10}$$

3. In Julia set,

$$C = \frac{\sum_{i=1}^{6} e_i C_i}{|\sum_{i=1}^{6} e_i C_i|} \tag{11}$$

   Because the elements of $\boldsymbol{\epsilon_2}$ is represented the emotion as mentioned in Section 2.3.2, C is calculated by considering the emotional intensity $\boldsymbol{\epsilon_2}$.
4. All pixel color of (x, y) is calculated as following $\boldsymbol{RGB}(r, g, b)$:

$$r = n_{\text{div}}(x, y)^3 \% v_1 + w_1 \tag{12}$$
$$g = n_{\text{div}}(x, y)^2 \% v_2 + w_2 \tag{13}$$
$$b = n_{\text{div}}(x, y) \% v_3 + w_3 \tag{14}$$

   In case of $\boldsymbol{\epsilon_1} = (0, 0)$,

$$r = g = b = n_{\text{div}}(x, y) \% 256 \tag{15}$$

   where % is a modular operator. $\boldsymbol{v} = (v_1, v_2, v_3)$ and $\boldsymbol{w} = (w_1, w_2, w_3)$ are defined as follows utilizing emotional vector $\boldsymbol{\epsilon_1} = (\epsilon_1, \epsilon_2)$. If $v_i = 0$, we reset $v_i = 1$.

$$\boldsymbol{v} = \left(\left[255\frac{\epsilon_2}{\epsilon_1 + \epsilon_2}\right], \left[255\frac{|\epsilon_1 - \epsilon_2|}{\epsilon_1 + \epsilon_2}\right], \left[255\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}\right]\right) \tag{16}$$

$$\boldsymbol{w} = (255 - v_1, 255 - v_2, 255 - v_3) \tag{17}$$

   where $[x]$ represents the maximum integer value which is not over x. In general, if the first element of $\boldsymbol{\epsilon_1}$ which represents positive is large, the main color of image becomes red. If the second element of $\boldsymbol{\epsilon_1}$ which represents negative is large, the main color of image becomes blue. If positive value and negative value are almost equivalent, the main color of image becomes green. If positive value and negative value are 0, the color of image becomes gray-scale.

Table 1. The Result of Questionnaire

label\fig.	fig.1	fig.2	fig.3	fig.4	fig.5	fig.6
1: joy	0	1	0	2	8	1
2: anger	4	3	2	1	0	2
3: sadness	3	0	0	1	2	6
4: hate	1	4	2	5	0	0
5: fear	4	3	1	1	0	3
6: surprise	0	1	7	2	2	0
emo. label	fear	anger	surprise	hate	joy	sadness

3 Assign Emotional Labels to Images

In this section, we show how to assign each emotion to images created by each $C_i (i = 1, 2, ...6)$ in 2.4.4. First of all, we have to describe that finding the strict relation between emotions and images is one of the most difficult problem, so we utilized imperfect rule to assign emotion labels in this paper. We focus on only 6 Julia set images by initial values C_1, C_2, ..., C_6, and assign only one emotion to each image by result of human questionnaire. Figures $1 \sim 6$ show Julia set images of $C_1, ...C_6$, and table 1 shows the result of questionnaire. The algorithm of assignment is as follows:

1. Set the total number of emotional labels j assigned to image i to x_{ij}. The number of emotion labels j is set to $\{1 \rightarrow$ joy, $2 \rightarrow$ anger, $3 \rightarrow$ sadness, $4 \rightarrow$ hate, $5 \rightarrow$ fear, $6 \rightarrow$ surprise$\}$. Where $\sum_{j=1}^{6} x_{ij} = 12$ and $\sum_{i=1}^{6} x_{ij} = 12$.
2. Denote $i,\ j$ of $\arg\max_{i,j} x_{i,j}$ to i^*, j^*
3. Assign emotional label j^* to image i^*, then remove all elements of x_{i^*j} and x_{ij^*} from $\{x_{ij}\}$.
4. If $\{x_{ij}\} \neq \emptyset$, go to step 3.

In this result, the reliability of label of the figure 5 (label : joy) and the figure 3 (label : surprise) is very high because the number of votes is large compare to other emotional labels.

We give one sheet printed gray-scale images to participants, so the order does not affect result. We obtained participants comment as follows:

- Because adding emotion labels to a certain gray-scale images is very difficult, lots of annotators adopted elimination way to annotate.
- The shape features of the image such as roundness and ruggedness has strong affect to results.

Table 2. Elements of *Kansei* Vector $\boldsymbol{\epsilon_2}$ of Each Input

emotion	input1	input2
joy	0.1169	0.1424
anger	0.0960	0.1186
sadness	0.3693	0.2568
hate	0.0440	0.0674
fear	0.0574	0.0827
surprise	0.1187	0.1673

4 Computer Experiment

To show the results of the proposed method mentioned in Section 2, we created 2 images based on 2 example inputs by the proposed method.

Input1 "I have a date with a pretty girl".

Input2 "I was really disappointed with a broken heart".

Figures 7 and 8 show the images created by **Input1** and **Input2** respectively. In figure 7, the value of element r in $\boldsymbol{RGB}$ is large because positive input. On the other hand, in figure 8, the value of element b in $\boldsymbol{RGB}$ is large because of negative input.

The number of morphemes[1] of **Input1** is 4, where the words of part of speech, auxiliary verbs, unknown words and symbols are removed. There exists 2 words (*pretty, date*) which have the positive sentiment in sentiment dictionary. On the other hand, there are no negative words in **Input1**. Therefore, $\boldsymbol{\epsilon_1}$ of **Input1** is calculated as follows.

$$\boldsymbol{\epsilon_1} = \left(\frac{2}{4}, 0\right) = (0.5, 0) \tag{18}$$

The number of morphemes of **Input2** is 5, where the words of part of speech, auxiliary verbs, unknown words and symbols are removed. There are no positive words in input2. On the other hand, there exists 2 words (*broken heart, disappointed*) which have the negative sentiment in the sentiment dictionary. Therefore, $\boldsymbol{\epsilon_1}$ of **Input2** is calculated as follows.

$$\boldsymbol{\epsilon_1} = \left(0, \frac{2}{5}\right) = (0, 0.4) \tag{19}$$

Table 2 shows $\boldsymbol{\epsilon_2}$ calculated by the algorithm in 2.3.2.

[1]We count the number of morphemes in Japanese

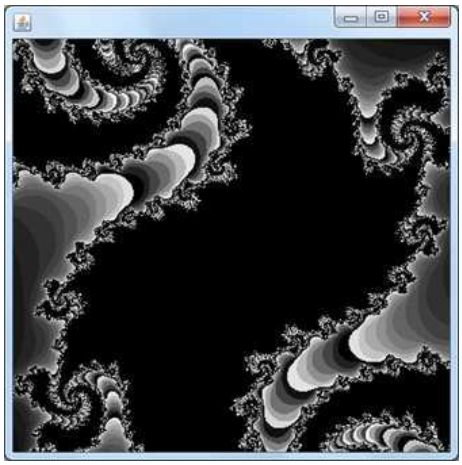


Figure 1. Image of the Julia Set : 1

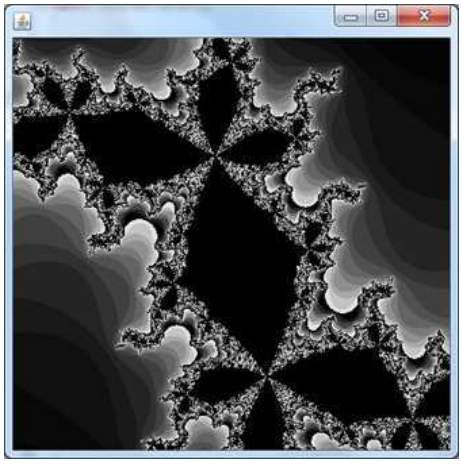


Figure 2. Image of the Julia Set : 2

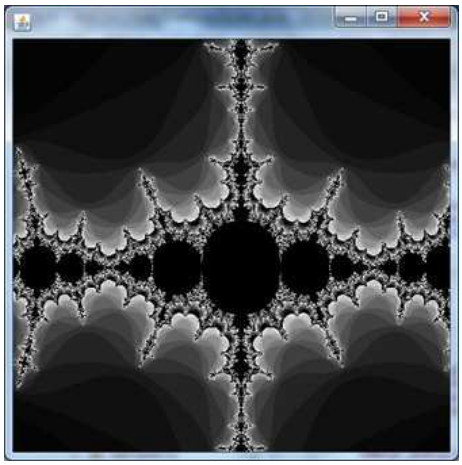


Figure 3. Image of the Julia Set : 3

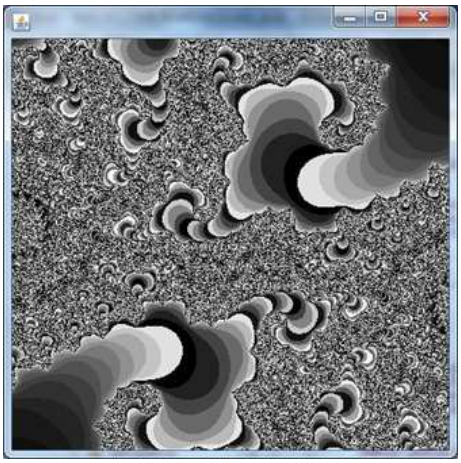


Figure 4. Image of the Julia Set : 4

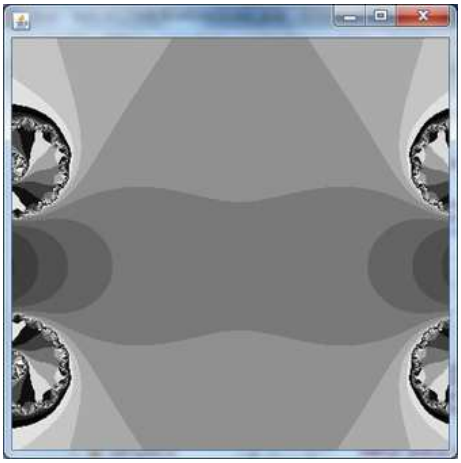


Figure 5. Image of the Julia Set : 5

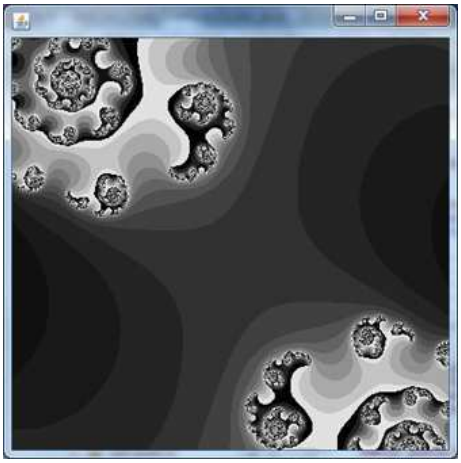


Figure 6. Image of the Julia Set : 6

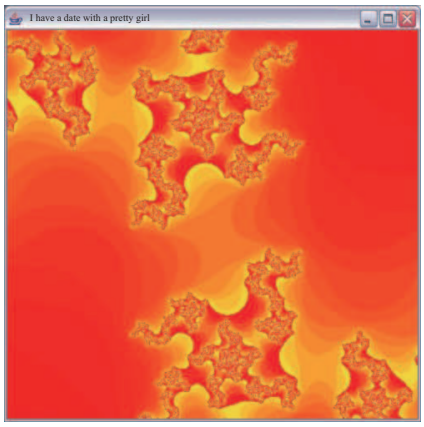


Figure 7. Image of the **Input1**

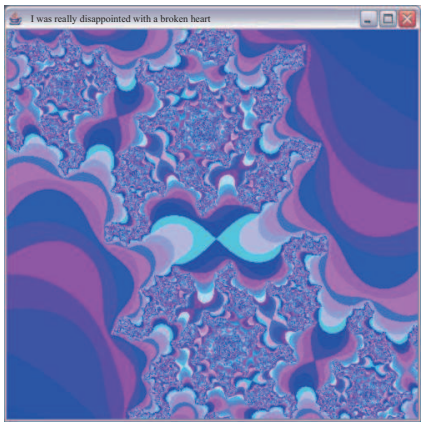


Figure 8. Image of the **Input2**

In this experiment, the problem is that several elements of ϵ_2 shown in table 2 do not match to human common sense because the intensity of sadness in **Input1** and the intensity of joy in **Input2** become large compare to the other elements. Improving algorithm of making ϵ_2 is the important future work.

5 Expand Dictionary

There is no theoretical way to define sentiment words vector χ even though χ is very important to calculate ϵ_2.

To define χ with reasonable grounds, we try to extend definition of χ to several candidate words which are related to 6 emotions and finding synonyms of these words referring to the thesaurus in order to create *emotional candidate words dictionary*. After making *emotional candidate words dictionary*, we find the best words for elements of χ by utilizing equation (1). We expect that the appropriate χ with best elements can represents the emotion of sentences correctly.

Table 3. The number of words in Emotional candidate words dictionary

emotion	words
joy	458
anger	177
sadness	243
hate	118
fear	181
surprise	107

Table 3 shows the number of words in each *emotional candidate words dictionary*. We try to add new words to *emotional candidate words dictionary* by machine learning with twitter results.

6 Application

In this section, we show the example of applying proposed method to an application.

We have researched picture drawing system by computer and have already proposed Twitter Bot called Cotopaint which can draw pictures semi-automatically according to user's tweets. Cotopaint generates a picture by combining prepared parts and simple figures such as circles and squares. We try to introduce the proposed method into the drawing algorithm of Cotopaint.

6.1 Outline of Cotopaint

The main procedures of drawing pictures in Cotopaint are checking user's tweet, executing selected drawing commands and uploading the picture with some comments to Twitter server. Generating pictures have 2 different layers called the base layer and the drawing layer. Base layer represents background of the picture such as blank or scenery pictures. Drawing layer represents variable parts of pictures and is changed everytime Cotopaint receives user's tweets.

Outline of Cotopaint is as follows:

1. Cotopaint tweets in order to inform users about starting to draw. Let $t = 0$. Set minimum waiting time m randomly in range [3,5]. Set variable waiting time c randomly in range [0,15].

2. Select base layer picture from the background picture database.

3. Wait $m + c$ minutes to receive user's tweets.

4. After $m + c$ minutes, get all tweets mentioned to Cotopaint to decide drawing commands. Execute commands to re-draw drawing layer.

5. Reply to each user in order to inform using drawing commands.

6. Output picture is generated by combining the base layer and drawing layer. This output picture of t is saved as new picture P_t into local disk.

   Then, upload new output picture to Twitter server. Set output picture of t to base layer of next time step $t+1$. Clear drawing layer.

7. Set c randomly in range [0,15]. Set $t = t + 1$. Go to step 3.

Figure 9 shows example of Cotopaint's Twitter timeline.

6.2 Drawing Command of Cotopaint

A drawing command $\boldsymbol{d}$ has 5 drawing parameters. $\boldsymbol{p}$, $\boldsymbol{c}$, n, r, s are defined as follows:

$\boldsymbol{p}$ **:** (x, y) coordinates of the drawing layer.

$\boldsymbol{c}$ **:** RGB color represented by (r, g, b).

n **:** Times of executing a drawing command.

r **:** The decided size for geometric pattern represented by pixel(px).

s **:** Type of drawing pattern.

Those parameters are decided based on user's inputs. Cotopaint recognizes specific keywords in inputs and translates those keywords into numerical drawing parameters. Cotopaint draws type s pattern, n times with the radius r, the color $\boldsymbol{c}$, the center location $\boldsymbol{p}$.

Table 4 shows relation between drawing parameters and keywords in user's input tweet. Parameter r has no keywords because r is decided randomly.

Cotopaint can execute following procedures by receiving user's mention tweets.

1. Draw figures by using the drawing command $\boldsymbol{d}$.

2. Set drawing parameters.

Table 4. Relationship between $\boldsymbol{d}$ and keywords

parameter	keyword	value
$\boldsymbol{p}$	"above"	$\boldsymbol{p} = (0, a)$
	"below"	$\boldsymbol{p} = (0, S/2) + (0, a)$
	"right"	$\boldsymbol{p} = (S/2, 0) + (a, 0)$
	"left"	$\boldsymbol{p} = (a, 0)$
	"center"	$\boldsymbol{p} = (S/2 - a, S/2 - a) + (a, a)$
$\boldsymbol{c}$	prepared color	set new color by the certain keyword
	others	set new color randomly
n	1 - 9 times	set n by the certain number
	"many", "lots of"	$n = 20$ (currently)
s	"circle"	set circle
	"square"	set square
	prepared parts	set parts by certain keywords

- S : The pixel size of the side of regular square canvas. S of Figure 9 is 400.
- a : $[0, L]$ uniform random number.
- L : The maximum range of a. L is decided by Cotopaint subject to $0 \leq L \leq S/2$.

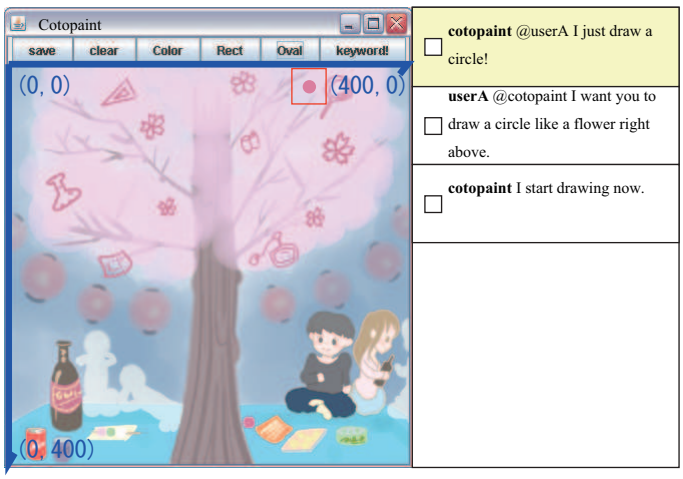


Figure 9. The example of timeline

3. Undo the previous commands.

4. Clear the drawing layer.

5. Save the base layer and the drawing layer.

6. Update the generated picture to Twitter server.

6.3 Problem of Cotopaint

Cotopaint needs user's instructions to draw a picture. Hence there are no surprise to see the generated picture. We would like to modify this feature of Cotopaint by introducing an unpredictability into generated pictures. However, we want to avoid showing the random pictures

and keep some relationship between user's inputs and generated pictures.

Proposed method in this paper may be useful to achieve this because generated fractal figures are hard to be predicted and those figures have some relationship such as polarity of positive/negative of user's inputs shown in 2.4.4. To analyze relationship between the generated picture and user's Kansei, we have to continue to research about quantitative representation of Kansei.

7 Conclusion

In this paper, we propose the novel method which produces images based on user's inputs and show experimental results.

Followings are subjects for further study.

Considering emotional vector In this paper, we try to represent humans emotions by the intensity of positive and negative, and the intensity of joy, anger, sadness, hate, fear and surprise. However, human emotions are more complex and consists lots of aspects. In order to represent human emotion in detail, we have to consider the definition of elements of emotional vector.

Creating various images In this paper, we utilize the Julia set as emotional image. However, there are lots of another types of images such as the Mandelbrot set, symbolic, graphic and cartoony images. Find an adequate relationship between emotions and images are important topic of this study.

Creating images with multi-labeled emotion Real emotion contains several basic emotions. For example, love is composed of joy, fear and another complex emotion. In order to translate all emotions to images, we have to show the reasonable algorithm of creating images which have relation to several emotions. However, it is difficult to generate figures which represent several emotions together. Our proposed method, a figure of mixing 2 emotions like joy and fear does not always have relation to fear and joy. Creating multi-labeled emotions figures is important further study.

Acknowledgment

This work was supported by Grant-in-Aid for JSPS Fellows(Grant Number:25·10978).

References

[1] Vilayanur S Ramachandran and Edward M Hubbard. Synaesthesia–a window into perception, thought and language. *Journal of consciousness studies*, 8(12):3–34, 2001.

[2] Daphne Maurer, Thanujeni Pathman, and Catherine J Mondloch. The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental science*, 9(3):316–322, 2006.

[3] Kazunari Ito, Motohiro Matsuda, MJ Dürst, and Koiti Hasida. Svg pictograms with natural language based and semantic information. In *Proceedings of the 5th International Conference on Scalable Vector Graphics (SVG Open 2007)*, 2007.

[4] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. Textual affect sensing for sociable and expressive online communication. In *Affective Computing and Intelligent Interaction*, pages 218–229. Springer, 2007.

[5] Michal Ptaszynski, Fumito Masui, Pawel Dybala, Rafal Rzepka, Kenji Araki. Open Source Affect Analysis System with Extensions. In *Proceedings of the 1st International Conference on Human-Agent Interaction(iHAI 2013)*, II-2-p9. 2013

[6] Hiroya Takamura, Takashi Inui, and Manabu Okumura. Extracting semantic orientations of words using spin model. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 133–140. Association for Computational Linguistics, 2005.

[7] Akihiko Yasui, Miki Ueno, Naoki Mori, Keinosuke Matsumoto. An Image Creating System based on Kansei of Natural Language Inputs, *Proceedings of the 57th Annual Conference of the Institute of Systems*, 337-2(CD-ROM), Control and Information Engineerings (ISCIE), 2013.

[8] Gaston Julia. Memoire sur l'iteration des fonctions rationnelles. In *Journal de Mathematiques Pures et Appliquees*, vol. 8, pages 47–245. 1918.

A Method to Classify Sentences of Topics and Opinion in Online Short Letters about Stock Trading Information Using Support Vector Machine

Makoto Okada **Kouhei Shimada** **Kiyota Hashimoto**

Osaka Prefecture University
{okada@mi.s, st301020@edu, hash@kis}.osakafu-u.ac.jp

Abstract

Recently, e-commerce has been flourishing and information of stock trading has been distributed using E-mail and other web contents via Internet. In this situation, traders need to analyze and select which information is important. It is important for the analysis of the information to distinguish between topic and opinion in the information. In this paper, we proposed a method to classify sentences of short letters about stock trading into "topic" or "opinion" using Support Vector Machine (SVM), one of machine learning method. We appended label of "topic" or "opinion" to the sentences of the short letters considering sentence-end expressions. The labeled data were used for learning data of SVM and test date. We estimated effectiveness of the proposed method by experimental results.

Keywords: Support Vector Machine (SVM), Sentence-End Expression, Stock Trading, Short Letter

1 Introduction

Recently, e-commerce has been flourishing. Under the situation, information of stock trading also has been offered in public on the network. And many people use some services that deliver the information of stock trading every day.

There are many studies using the stock trading information. Hirokawa et al. [1] analyzed relationships between trends of stock prices and articles of newspapers, and difference of power of influence of words / phrases in articles of newspapers on stock prices in each type of business.

Short letter of stock price is one of text based information. It sends various news articles related to the stock price to readers. The readers can use the information contained of the short letter to their stock trading.

Generally, sentences contained the short letter categorized two types. One is topic, another is an opinion. An opinion sentence is described by a author based on his / her conjectures and / or estimates. For readers that read the short letter, it is important to show the sentences with distinction of their types. Based on the distinctions, the readers can decide stock trade appropriately without confusion of information of topics and opinions and they also avoid a loss.

And when there are several opinions to one topic, it is useful for the readers to show the opinions that are extracted from the short letters automatically. It makes readers be able to compare the opinions and decide appropriately.

In this paper, we will propose a method to classify sentences of the short letters as topic and opinion automatically. We consider the proposed method as a preliminary step in order to extract information of topics and opinions automatically. The proposed method use support vector machine (SVM) which is a method of machine learning.

In section 2, we will explain an outline of a short letter about stock information and detail of sentence of topic and opinion. In section 3, we will explain an outline of support vector machine (SVM). In section 4, we will show experimental results and discussions. In section 5, we will summarize this paper and describe future works.

2 The short letter of stock information, sentence of topic and opinion

The short letter of stock information is a text style news data which is sent to peoples who trade stocks by e-mail. In this study, we used the short letter data written by Japanese. The shortest data of one short letter contained about 20 characters and the longest one contained over 1,000 characters. The average of the number of characters in one short letter is 803 characters.

2.1 Sentences of topic and opinion

Generally, the short letter contains two types of sentence, topic and opinion. Sentences of topics describe facts and news related stock trading. On the other hand, sentences of opinions express opinions and expectations of authors of the short letters based on these topics.

In this study, first, we selected 1,000 short letters and extracted 7,440 sentences contained in the short letters. We appended labels of topic or opinion to the sentences by human before experiments. Table 1 shows the data about numbers of sentences of each label. Each row show numbers of sentences in every 100 short letters. Total sentences of topic were 6,593 and others of opinion were 847.

Table 1: Number of sentences of topic and opinion.

No. of Set of Short Letter	Number of Sentence of "Opinion"	Number of Sentence of "Topic"
1	93	630
2	97	541
3	96	804
4	98	713
5	36	596
6	112	823
7	25	377
8	127	890
9	70	403
10	93	816
Total	847	6593

2.2 Criterion of label appending

In this study, we set a criterion for appending labels to sentences. If some sentence satisfied the criterion, we append a label "opinion". Other sentences were appended a label "topic".

In this study, we devise some sentence-end expressions can be used as a criterion to decide whether a sentence describes "topic" or "opinion". We selected some sentence-end expressions shown by Figure.1 for decision of sentence. If these expressions are contained in a sentence, the sentence was decided "opinion". If they are not contained, the sentence was decided "topic".

Figure 1. Sentence-end exprassions contained the sentences of opinion.

seems <to be, to do>

- ～もよう ("Moyoh"),
- ～よう(だ)("You(da)")
- ～そう(だ)("Sou(da)")
- ～とみられる("To-Mirareru")

want to do

- ～しておきたい("Shite-Okitai")
- ～したい("Sitai")

may <be, do>

- ～かもしれない("Kamo-Shirenai")

<suppose, hope, is afraid> that

- ～だろう("Darou")

Generally, the short letters contain two types of sentences, topics and opinions, and mostly they do not contain other types of sentences. And most sentences contained them describe about topics. Sentences about opinions in them are very few.

Therefore, we can estimate which label is appended to a sentence, “topic” or “opinion”.

When learning data and test data of SVM are constructed from the sentences in the short letters, sentence-end expressions are used as a criteria of “Opinion” and “Topic”. However, when SVM constructs a classifier using the learning data, every texts, also the sentence-end expressions, are divided to morphemes. Therefore, the sentence-end expressions are not used directly in learning phase of SVM.

3 Support Vector Machine (SVM)

A support vector machine (SVM) is a supervised learning method for binary classification, originally proposed by Vapnik [2]. Figure 2 shows the basic concept of SVM.

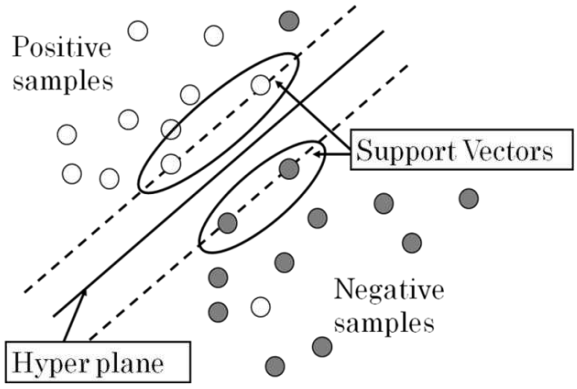


Figure 2. An example of SVM

The algorithm computes the hyperplane which divides two classes by analyzing the vector data of two classes as teacher data. With the hyperplane, the machine estimates the class of unknown data. Even if vector is high dimensional, the machine can analyze. Since the dimension of vector of documents data is often high, the machine is used for analyzing of natural language.

In this study, we combined kernel function with SVM. The kernel function enables to calculate inner product in feature space from data directly without calculation using coordinates of data. Therefore, to use kernel function is useful to decrease total amount of calculation of inner product and to reduce calculating time.

4 Experiments and Discussions

4.1 Experiments

First, we constructed learning and test data from short letters. We appended label “topic” or “opinion” to 7,440 sentences in 1,000 short letters of stock information shown by human (Table.1). We use these data as both of learning data of SVM and test data for experiments. The learning data did not overlap the test data.

We evaluated effectiveness of SVM for classification “topic” and “opinion” under several conditions. In this study, we used LIBSVM library module and kernel function is Gaussian (RDF) kernel.

We use three values, Precision(P), Recall(R) and F-measure(F-M) in order to measure effectiveness of classification of SVM in each experiment. We use 10-fold cross validation in every experiment.

In experiment 1, we use all data shown on Table.1. 1,000 short letters were divided into 10 units. Each unit contained 100 letters. 900 letters were chose to learning data of SVM and remaining 100 data used test data. Every sentence in all of the letters were divided to morphemes using MeCab which was Japanese Part-Of-Speech (POS) and morphological analysis program. After the analysis, SVM learned these results and constructed a classifier and the test data classified by the classifier. When the classifier were constructed, the sentence-end expressions separated into morphemes, therefore the expressions were not used special triggers for classification in this study.

Next, we investigated an effect of using same number of sentences of topic and opinion when SVM made a classifier by learning the data. We chose 1,400 sentences, a half of them is topic, and another half is opinions. And we found that there were same sentences in different short letters because same news was distributed by different letters in several times. Therefore, some sen-

tences were overlapped. Because of it, we made two data set, one left overlapped sentences and another removed overlapped sentences. We also investigated effect of the difference of the overlapped sentences to learning and classification. Table 3 shows the experimental results

Final, we investigated which combination of Part-Of-Speech (POS) was effective for classification using SVM by experiment. Figure 3 shows the combinations of POS using the experiment. Experimental data was same as the data of Experiment 2. In experiment 3, several classifiers were constructed by SVM using morphemes that satisfied the combination of POS. And we used these classifiers to the experiment 3. Table 4 shows results of the experiment 3.

4.2 Discussions

In the results of the experimental 1, the precision showed high ratio both of "topic" and "opinion". However, the recall and the F-measure of "opinion" were lower than the value of "topic".

Table 2: Results of experiment 1

Short Letter	Sentence "Opinion"			Sentence "Topic"			
	P	R	F-M	P	R	F-M	Avg. F-M
1	87.7	53.8	66.7	93.5	98.9	96.1	81.4
2	100	57.7	73.2	93.0	100	96.3	84.8
3	100	64.6	78.5	95.9	100	97.9	88.2
4	100	67.4	80.5	95.7	100	97.8	89.2
5	100	33.3	50.0	96.1	100	98.0	74.0
6	95.7	59.8	73.6	94.8	99.6	97.2	85.4
7	100	36.0	52.9	95.9	100	97.9	75.4
8	100	56.7	72.4	94.2	100	97.0	84.7
9	96.0	68.6	80.0	94.8	99.5	97.1	88.5
10	50.0	2.2	4.1	89.9	99.8	94.6	49.4
Avg.	96.9	52.4	68.0	94.2	99.8	96.9	82.5

Table 3: Results of experiment 2 (same number)

F-Measure	Sentence "Opinion"	Sentence "Topic"	Average
Experiment 1	68.0	96.9	82.5
Experiment 2 Overlapped Data	82.4	80.0	81.2
Experiment 2 Not-Overlapped Data	79.3	72.3	75.8

(a) All,

(b) Verb,

(c) Auxiliary Verb,

(d) Noun,

(e) Noun, Postpositional Particle,

(f) Verb, Auxiliary Verb, Noun,

(g) Verb, Auxiliary Verb, Postpositional Particle,

(h) Verb, Auxiliary Verb, Noun, Postpositional Particle

Figure 3. Combinations of POS.

Table.4: Experimental results (same number, combination of POS)

	Results of Overlapped Data						Results of Non-Overlapped Data					
	Sentence of "Opinion"			Sentence of "Topic"			Sentence of "Opinion"			Sentence of "Topic"		
	P(%)	R(%)	F-M	P(%)	R(%)	F-M	P(%)	R(%)	F-M	P(%)	R(%)	F-M
(a)	77.8	87.6	82.4	85.8	75.0	80.0	70.4	90.7	79.3	87.0	61.9	72.3
(b)	73.6	76.4	75.0	75.5	72.6	74.0	69.5	85.7	76.8	81.4	62.4	70.6
(c)	87.8	64.7	74.5	72.1	91.0	80.5	85.5	72.4	78.4	76.1	87.7	81.5
(d)	91.8	62.6	74.4	71.6	94.4	81.4	85.4	61.2	71.3	69.7	89.6	78.4
(e)	94.5	68.7	79.6	75.4	96.0	84.5	91.3	64.3	75.5	72.4	93.9	81.8
(f)	95.9	73.7	83.3	78.6	96.9	86.8	83.9	84.0	83.9	84.0	83.9	83.9
(g)	76.1	93.3	83.8	91.3	70.7	79.7	75.1	96.2	84.4	94.7	68.2	79.3
(h)	82.8	83.7	83.2	83.5	82.6	83.0	72.0	90.1	80.0	86.8	64.9	74.3

Number of sentences of "topic" were more than number of sentences of "opinion". We suppose the imbalance of the data gave bad influence to the recall and the F-measure.

In experiment 2, we made number of sentences of "topic" and "opinion" even and values of the F-measure of "topic" were almost equal to the F-measure of "opinion". We confirmed that learning balanced data will improve performance of classifier by SVM.

We also found influence of overlapped data by the value of F-measure. The value of F-measure used overlapped data were higher than the value of F-measure used no-overlapped data. We can estimate that the overlapped data emphasized characteristics of "topic" and "opinion".

In the experiment 3, the value of F-measure make it possible to estimate that the combination of (Verb, Auxiliary Verb, Postpositional Particle) is effective for classifying the sentences of topic. The sentence-end expressions are constructed by these POS. The result shows the sentence-end expression are effective for classification of "Topic" and "Opinion". The sentences of topic were classified adequately by learning the combination containing "Noun". In this experiment, the best combination of POS is (Verb, Auxiliary Verb, Noun).

5 Conclusion and Future Works

In this paper, we proposed a method to classify sentences of “topic” and “opinion” of the short letters of stock information by the classifier by SVM. We confirmed effectiveness of the proposed method by the experiments.

By the experimental results, we confirmed the effectiveness of making numbers of sentences of “topic” and “opinion” even in order to construct good classifier by SVM. We also confirmed the overlapped learning data have the tendency to emphasize the characteristics. And we can improve an accuracy of classification by considering POS information of the learning data.

Future works are as follows. (1) other criterions between “topic” and “opinion” except sentence-end expressions [3], (2) investigation of increase of number of data, (3) comparison the results of other machine learning methods.

References

[1] Takamasa Hirokawa, Minoru Yoshida, Hidetaka Masuda, Hiroshi Nakagawa. An Analysis of Relationships between Trend of Stock Price and News Articles considering Industrial Classification (Japanese). *The 16th Annual Meeting of the Assosiation for Natural Laguage Processing*, 1070-1073, 2010.

[2] V. N. Vapnik. The Nature of Statistical Learning Theory, 2nd ed. *Springer-Verlag*, New York, 2000.

[3] Toshihiro Kawaguchi, Tohgoroh Matsui, Hayato Ohwada. Opinion Extraction from Weblog using SVM and Newspaper Article. *The 20th Annual Conference of the Japanese Society for Artificial Intelligence*, 1A3-3, 4 pages, 2006.

Topic Sentence Extraction from Editorial Articles Based on Sentence Structure and Topic Relevance

Muneaki Harada† **Hidekazu Yanagimoto**† **Michifumi Yoshioka**†
†Osaka Prefecture University
{harada@sig.cs, hidekazu@kis, yoshioka@cs}.osakafu-u.ac.jp

Abstract

Today there are quite a lot of articles not only on published media but on the Internet. Many people represent their own opinions related to political and economic problems. However, it is some troublesome to read all their articles and understand author's main statement from each article since the number of articles is so large. The purpose of this study is extracting topic sentence using three approaches. The first one considers mood which is writer's tone and attitude. We examine what word and expression increase sentence importance with multiple regression analysis. The second approach is considering topic relevance with Latent Dirichlet Allocation. A paragraph which contains words related to article's main topic is more important than other paragraphs including less relevant words. We emphasize paragraphs which highly related to the main topics. The third approach is focusing on rhetoric. A sentence like " Not A but B " includes author's strong opinion. Finally summing score of each sentence score based on these approaches, we find sentences representing author's opinion clearly.

Keywords: Natural Language Processing, Summarization, Regression Analysis,LDA

1 Introduction

For many kinds of editorial articles, many analysts and writers insist their opinion in their own blogs and columns. However, it is difficult to grasp the gist of their argument enough because of the quantity of articles. The aim of this study is extracting important sentences which represent author's opinion clearly. There are some ways for determining the importance of sentence, such as the weight of words, the position of sentences, structure and expression. Hirao[1] says it is important to combine these clues effectively.

SumBasic [2] is a good approach for extracting important sentences. This method uses word frequency only. Hence, there is room for improvement. Yie [3] uses not only frequency but also position information. However, documents which have important sentences in the beginning of documents are not so much. Position information depends on a type of documents deeply. Furthermore, we need to think other clues which make a sentence more important. For solving these problems, we propose approach using three methods. First we focus on mood, which represents author's tone and attitude. In Japanese there are various kinds of meaning at the end of a sentence, which can represent not only tense but also confidence, tone, regard, and so on. In traditional study what word has importance is determined by hand. However, it tends to depend on one's sense and is a laborious task. In this study we use multiple regression analysis and determine the score of every expression of a sentence. Then we estimate the importance of each sentence by using the score of expression. Second, we use the topic with LDA for evaluating the relation between a whole document and each paragraph. A paragraph which has little relevance with article's topic is not useful; therefore, we calculate how much topic relevance each sentence has. Finally, we focus on rhetoric and determine the final score of every sentence.

2 Proposed Method

We usually find the gist of articles based on background knowledge. However, it is difficult to explain why we select certain sentences to summarize them. Machine learning needs correct labels of input data or rules to recognize the importance. In this study we estimate the importance of each sentence using mood, topic relevance and rhetoric. Using these approaches, we calculate how important each sentence is and evaluate how much each sentence describes what a writer wants to say.

2.1 Importance Prediction Considering Mood

In editorial article, a sentence which describes opinion is an important sentence. You find out these sentences based on background knowledge. However, it is difficult to make strict rules. First of all, we focus on expression called mood. Mood is one of grammatical forms which represent author's tone and attitude. The words representing mood often exist on the beginning of a sentence, the end of a sentence or the end of a clause. We think each mood affects the importance of a sentences and it is useful for estimating the importance of sentence. We prepare training data labeled with 5 grades by evaluators and determine mood expressions with multiple regression analysis.

The expression representing mood is at the beginning of a sentence, the end of a sentence or the end of a clause. In this study we use part of speech to extract these expressions at first and determine the strength of impression of these words. To determine part of speech, we use MeCab[4] - morphological analysis tool and CaboCha[5] - dependency analysis tool.

The number of each expression words is explanatory variable and we add the following elements into explanatory variable.

- the number of adnominal adjectives
- the number of question marks
- the number of brackets
- the number of proper nouns
- the frequency of proper nouns
- the number of clauses

Let Y be defined as

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_i + \gamma \quad (1)$$

where Y is the importance of a sentence, α is intercept, β_i is coefficient of each words, X_i is frequency of words. γ represents influence of importance of a previous sentence since we assume that sentences relate to each other and the importance of sentences don't change quickly. Training data is made by hand and labeled with 5 grades.

2.2 Importance Prediction Considering Topic Relevance

After analyzing each sentence considering mood, we can find sentences which describe author's opinion strongly. However, these extracted sentences should be not only strong but also well relevant to the topic of an article. Therefore, we predict a topic composition of an article and each paragraph, emphasizing the paragraph well related to the main topic. In this study we use Latent Dirichlet Allocation (LDA)[6] for topic analysis. LDA is one of topic models assuming that one article consists of various topics and each topic is formed by a set of words having a generated probability. Learning is carried out with Collapsed Gibbs Sampling. We calculate the topic vectors whose elements are words generating probability of each topic with LDA and compute topic similarity between each paragraph and article with cosine similarity. Topic vector is normalized for comparing.

2.3 Importance Prediction Considering Sentence Structure

Just using words representing strong mood do not mean they are important. Authors would pay attention to sentence structure for better understanding. In this study we focus on rhetoric and consider further expression in sentences representing opinion. First we assign a role to every sentence. We categorized every sentence into four roles, " Fact ", " Neutral ", " Opinion " and " Question " based on heuristic rules considering the expression at the end of sentence and the word meanings. For example, past tense is thought to be fact, while a sentence including " think ", " should " or " want to " is thought to be opinion.

Opinion following fact sentences is thought to represent author's opinion clearly. This opinion is based on facts and the first sentence of opinion sentences is more important than others. We weight the opinion sentence after fact sentences. Opinion sentence just before fact sentences means the last sentence before change of the subject. It is thought to be the small summary of opinion sentences. We think it becomes clue of importance and weight this sentence. Opinion just behind a question is thought to be an answer against the question. Asking oneself may be what the author wants to say strongly. We weight these sentences locating just behind ques-

tion. Opinion just behind a negative sentence like “ Not A but B ” is thought to emphasize author’s opinion using anti-examples. It is thought to be important and weight this sentence. The negative sentence is judged by including NOT in the sentence.

3 Experiment

We explain data in our experiments. We use 33 editorial articles appearing in BLOGOS[7] and AGORA[8] for generating topic model with LDA and we use 1806 articles in The Nikkei. We extract nouns in these articles and these words become candidates for generating a topic model. The 50 highest probability words in each topic are used for representing categories. The procedure for experiment is as follows.

1. After evaluating the importance according to 5 grades for every sentence by hand, we extract effectual expression words representing mood from articles and calculating the score of these words with morphological analysis. Using these score, we calculate the score of important expressions.

2. Generating topic model with LDA for 1,806 articles, we calculate generating probability of each topic and the topic vector of document and each paragraph.

3. Comparing the topic of document with each paragraph, we examine topic relevance of every sentence.

4. If there is particular sentence structure, we add score to the relating sentence.

5. Using 3 scores, we calculate the final score of each sentence. After considering sequence of sentences, we extract high importance sentences.

We show the top and bottom 10 words about the importance of every word in Table1, Table2. We were able to discover that words which refer to something such as “ koso ” (meaning emphatic phrase), “ nikanshite ” (as to), “ niyori ” (by), “ nitotte ” (for) increase the importance. Furthermore, we found that words which remaking something like “shika”(only), “koso” (That’s just) contribute to the importance. On the other hand the word presuming something like “areba” (If there is), “tatoeba” (for example) or weak tone words like “amari” (not so), “honno” (mare) lower the importance.

Table 1. Top 10 words weight

Expression	(English)	Score
koso	That’s just	1.282586
nikanshite	as to	0.609464
niyori	by	0.574194
nitotte	for	0.549843
dokoroka	not only	0.542623
nagara	while	0.505034
niwokeru	for	0.496711
dano	such as	0.465953
shika	only	0.415071
niyotte	by	0.384637

Table 2. Bottom 10 words weight

Expression	(English)	Score
totomoni	with	-0.539573
tsutsu	as	-0.532852
donnna	whatever	-0.469398
areba	if there is	-0.437094
amari	not so	-0.411040
tomo	at most	-0.388508
tatoeba	for example	-0.377526
nitsuite	about	-0.301774
honno	mare	-0.286686
oru	exist	-0.230216

3.1 Evaluation

Comparing three approaches, we confirm the efficiency.

- Conventional approach (depending on words frequency only)

- Proposed approach (using 3 methods)

- Hybrid approach (using 3 methods + words frequency)

Test data is as follows.

1. “ No blind spot with Abenomics ”
   - Yutaka Harada[9]

2. “ Abe’s exhausted policy ”
   - Nobuo Ikeda[10]

3. “ Operate inflation target flexibly ”
   - The Nikkei - Editorial[11]

4. “ Countermeasures for aging society ”
   - Tetsuzo Matsumoto[12]

5. “ Deteriorate of the media ”
   - Tatsuru Uchida[13]

6 persons read the original articles and extracted sentences. Using manual labeled data, we evaluate precision. Criterions for judgment are as follows.

1. Criterion1: How much extracted sentences represent author's opinion without omission.

2. Criterion2: How much extracted sentences relate to article's theme and don't include noises.

3. Criterion3: How much we can understand what authors want to say.

3.2 Examination

While the proposed method is inferior to the previous method in article 1 and 2, the hybrid method got the best results. These two articles are described about financial policy and require much knowledge on existing policy and the policy which has been carried out. Hence, it is difficult for those who don't have enough knowledge to understand articles. In this case it is easy to read and understand if extracted sentences include some facts and background knowledge on the theme. I think it contributes to good performance of the hybrid method. On the other hand the proposed method had good result in article 3, which is easy to read since it is editorial article on newspaper. Therefore it's easy to understand the extracted sentences which contain opinions only. Since Article 4 and Article 5 don't require special knowledge to read them, we got good results in the proposed method. Those articles consist of not only specific concern but also story of author's experience and episodes. This makes the frequency of each word diverse and it becomes difficult to grasp the main topic of the article precisely. On the other hand our proposed method focusing on just opinion is effective since it can ignore sentences which don't contain what an author really wants to say. If we focus on readability and summarization, we need to reconsider how much fact should be included in extracted sentences.

Table 3. Result

		Conventional	Proposed	Hybrid
Data1	crt1	3.0	2.7	3.7
	crt2	3.5	2.7	3.5
	crt3	3.3	2.7	3.2
	Total	9.8	8.0	10.3
Data2	crt1	2.7	2.2	3.0
	crt2	2.8	2.7	3.7
	crt3	2.2	1.7	3.0
	Total	7.7	6.5	9.7
Data3	crt1	2.8	4.5	3.5
	crt2	2.8	4.2	3.5
	crt3	2.2	4.8	3.8
	Total	7.8	13.5	10.8
Data4	crt1	3.0	3.5	3.7
	crt2	3.3	3.5	3.2
	crt3	3.3	3.5	3.7
	Total	9.7	10.5	10.5
Data5	crt1	2.8	3.8	3.8
	crt2	2.8	3.7	3.0
	crt3	2.8	4.2	3.5
	Total	8.5	11.7	10.3
	Score	43.5	50.2	51.7

4 Conclusion

Our method can achieve better sentence extraction comparing with the previous method by the combination of three approaches. However, there are some articles which got a better result by the proposed method and room of improvement. There are many other elements for increasing importance such as unknown information, fresh news and useful for readers. Considering these elements, we will try to improve the accuracy more.

References

[1] Tsutomu HIRAO, Hideki ISOZAKI, Eisaku MAEDA, Yuji MATSUMOTO, Important Sentence Extraction Based on Support Vector Machines 2003.

[2] A. Nenkova and L. Vanderwende.The impact of frequency on summarization.Technical report, MSR-TR-2005-101, 2005.

[3] Wen-tau Yih and Joshua Goodman and Lucy Vanderwende and Hisami Suzuki. Multi-Document Summarization by Maximizing Informative Content-Words, 2007.

[4] Taku Kudo, http://mecab.googlecode.com/

Table 4. Example of extraction with proposed method (Compressibility:28.7%)

The policy board's move is a down payment on Kuroda's pledge to bring an end to Japan's deflation through "monetary easing of far higher quality and quantity" than in the past. Re-energizing the economy will require making the most of the BOJ's actions, but the central bank also needs to pay close attention to possible side effects. There is no doubt that bold monetary easing is vital to pull the country out of a deflationary hole a decade or more in the making. The BOJ has good reason to take a more aggressive stance in trying to change market sentiment and improve the health of the real economy. But it must also be careful not to overdose on monetary stimulus. If the BOJ's new policy is viewed as an attempt to monetize the government's debt, the credibility of Japan's economic management will be seriously undermined. To avoid that, the government needs to demonstrate its commitment to improving Japan's fiscal health by raising the consumption tax and reforming the social security system. (The Nikkei - Editorial(April 5, 2013): "BOJ Move Welcome, But Only First Step"[14])

[5] Taku Kudo, Yuji Matsumoto, http://code.google.com/p/cabocha/

[6] David M. Blei, Andrew Y. Ng, Michael I. Jordan, Latent Dirichlet Allocation

[7] http://blogos.com/

[8] http://www.agora-web.jp/

[9] Yasushi Harada, FACTA - https://facta.co.jp/

[10] Nobuo Ikeda, NEWSWEEK http://www.newsweekjapan.jp/

[11] The Nikkei Editorial January 23, 2013

[12] Tetsuzo Matsumoto, AGORA - http://agora-web.jp/

[13] Tatsuru Uchida, http://blog.tatsuru.com/

[14] The Nikkei Editorial April 5, 2013

Table 5. Original of above extraction

The Bank of Japan decided Thursday to step up its quantitative easing program in hopes of freeing the econmy from the grip of deflation.

At its first policy board meeting under its new governor, Haruhiko Kuroda, the central bank moved to double the amount of money flowing in the economy in two years by increasing purchases of long-term government bonds and risk assets such as exchange-traded funds. The measures are aimed at reaching the BOJ's 2% inflation target within about two years.

The policy board's move is a down payment on Kuroda's pledge to bring an end to Japan's deflation through "monetary easing of far higher quality and quantity" than in the past.

Re-energizing the economy will require making the most of the BOJ's actions, but the central bank also needs to pay close attention to possible side effects.
It should remain flexible in implementing the new policy to avoid overshooting the inflation target, and be disciplined in its purchases of government bonds so that the measure is not seen as an attempt to plug Japan's huge budget hole.

According to the plan announced Thursday, the BOJ will double the amount of long-term government bonds on its balance sheet within two years. It will also simplify the purchasing method and increase its holdings of bonds with many years to maturity. And the bank will buy more of such risky assets as ETFs and real estate investment trusts. The BOJ said it will keep the expanded quantitative easing program in place until the annual rate of inflation rises to 2%, a target it hopes to hit within about two years.

Japan has had stream of good economic news in recent weeks. Prime Minister Shinzo Abe's expansionary economic policy has triggered a strong stock market rally and pushed the yen lower. Business and consumer confidence is showing signs of perking up. But there is no evidence of significant improvement of the real economy yet, with many economists still cautious about the economic outlook.

There is no doubt that bold monetary easing is vital to pull the country out of a deflationary hole a decade or more in the making. The BOJ has good reason to take a more aggressive stance in trying to change market sentiment and improve the health of the real economy.

But it must also be careful not to overdose on monetary stimulus. Too-aggressive monetary expansion could cause economic and market turbulence by pulling in "hot money" quickly, only to see it vanish should interest rates start to rise, for example. Instead of mechanically pumping up the money supply to reach the inflation target, the BOJ should adjust its operations to changing circumstances.

And monetary easing alone will not lead the economy back to the path of healthy growth. The government must do its part by promoting freer trade, loosening cumbersome regulations and improving the business environment through corporate tax cuts, among and other measures.

The BOJ's expanded program means its holdings of long-term government bonds will keep growing by around 50 trillion yen every year. That is far more than the amount of new government bonds to be issued every year under the government's initial budgets.
If the BOJ's new policy is viewed as an attempt to monetize the government's debt, the credibility of Japan's economic management will be seriously undermined.

To avoid that, the government needs to demonstrate its commitment to improving Japan's fiscal health by raising the consumption tax and reforming the social security system.

(The Nikkei - Editorial(April 5, 2013): "BOJ Move Welcome, But Only First Step"[14])

Exploiting WordNet for Wikipedia-Based Named Entity Disambiguation

Thuan K. Nguyen†* **Tru H. Cao‡***

† University of Science – VNUHCM
‡ University of Technology – VNUHCM
* John von Neumann Institute – VNUHCM
thuan.nguyen@jvn.edu.vn, tru@cse.hcmut.edu.vn

Abstract

Entity disambiguation is an important problem in semantic analysis and natural language processing. In this paper, we propose an approach to employ features of the WordNet ontology in the task of disambiguating named entities to Wikipedia. Methods of enriching text with synonymous relations of words are explored. An analysis of the results from our experiments shows that the accuracy of the disambiguation process is improved on a dataset of named entities.

Keywords: Entity Linking, Word Sense Disambiguation, Synset, Semantic Annotation

1 Introduction

The rapid growth of World Wide Web has been exposing an abundant and constantly updated data source to users around the globe. However, due to the unstructured nature of this large data collection; a crucial task of Natural Language Processing (NLP) is to better understand the content of a document through only some important keywords or phrases in the document. Such keywords or phrases are called entities and can be divided into two types: named entities (such as people, organizations and locations) and abstract entities expressing general concepts. Named entities carry essential information to convey the content of a document and thus are the central subject in many recent studies.

A named entity can be mentioned, or referred to, in a text by multiple possible surface forms (case 1) and conversely, a surface form can refer to multiple possible candidate entities (case 2). The latter case is where we need to perform disambiguation in order to tell which entity is actually referred to. The ability to identify named entities and properly link them to a knowledge base (KB) has been established as an important task in several areas, including topic detection and tracking, machine translation, and information retrieval [1, 4, 5, 10, 11]. This process is well-known as Entity Linking [9] or Entity Disambiguation [13]. Its main goal is to map mentions of entities into their correct referent entities in a provided KB used as a candidate inventory.

An example of case 1 described above is when "*the United States of America*" is referred to as "*the USA*", "*the U.S.*", or "*the States*". For case 2, an example is when we need to perform disambiguation in order to determine the mention "*Texas*" in a document refers to Texas as a state of the USA, or Texas as a Charlie Daniels single from the album Nightrider, etc… The output of the disambiguation process is the identifier (ID) of the KB entry to which the mention refers; or NIL if there is no such KB entry [9].

Thanks to its broad coverage of subjects as well as its well organized structure, Wikipedia has been widely used in NLP researches and applications. It can serve as a reliable ontology for the entity disambiguation problem [1] and hence is chosen as the candidate inventory in our approach. The disambiguation result is a corresponding Wikipedia article for each given mention in the text document, or NIL if there is no such article.

Previous works on named entity disambiguation have employed different approaches, including machine learning [4, 5, 11], vector space model and rule-based [1, 2, 3, 8] or graph-based [15]. Even though these approaches exploited diverse features and employed different models in named entity disambiguation, the lexical aspect of words in the text has not been explored. Based on the observation that a word sense might be represented by multiple surface forms in a

given context, we present a disambiguation approach that extends the one proposed in [2] by integrating WordNet features into the entity linking process. This idea, however, is not limited to our presented model and can be applied to other approaches.

In particular, by taking into account the synonymous relations of words, we expand the bag of words for a mention in the text with synsets of the words contained inside. The synset expansion is applied either to all words after a word sense disambiguation process, or to words that have only a single sense.

The rest of the paper is organized as follows. Section 2 presents a brief overview of related works on named entity disambiguation. Details of the WIN algorithm upon which our system is built as well as relevant concepts in WordNet are explained in Section 3.1 and Section 3.2, respectively, followed by our proposed approach in Section 3.3. Next, in Section 4, experiment results and analyses are presented. In section 5, we draw conclusions on the explored methods and discuss possible future works.

2 Related Works

A typical approach to named entity disambiguation involves the measuring of context compatibility between a mention and its candidate entities. The Wikify! system presented in [5] trains a Bayes classifier for each ambiguous mention using three words to the left and the right of outgoing links in Wikipedia articles, with their parts-of-speech, as contextual features. However, this method requires a large amount of processing effort because of the parsing of entire Wikipedia.

Bunescu and Paşca [8] propose a method that uses an SVM kernel to compare the lexical context around the ambiguous mention to that of its candidate entities, in combination with estimating correlation of the contextual words with the categories of the candidate entities.

Cucerzan [1] was the first to model interdependence among disambiguation decisions. In [1], the authors exploit semantic relatedness which is based on overlap in categories of entities referred to in the text.

Han and Sun [15] propose a method based on a referent graph that performs collective inference of KB entities referred to in a text. A referent graph is a weighted and undirected graph $G = (E, V)$ where V contains all mentions in the text and all possible candidates of these mentions. A mention-entity edge is established between a mention and an entity, and weighted based on context similarity, or a combination of popularity and context similarity. An entity-entity edge is established between two entities and weighted using semantic relatedness between them. The local context similarity is calculated based on a bag of words model and the semantic relatedness is computed using the formula presented in [4].

In [2], the authors present an incremental approach where already disambiguated entities help in the process of disambiguating other remaining entities. The incremental mechanism of this approach is similar to human processing of mentions based on previously known ones. This approach is also a novel hybrid approach of combining heuristic rules and statistics compared to other methods discussed earlier.

Although these existing works have used different strategies for entity disambiguation, only the exact forms of words are considered. The measuring of the context compatibility between a mention and its candidate entities thus has yet considered the case of multiple surface forms for a word. Consequently, our approach aims to explore the effect of using additional surface forms for a word in measuring the context compatibility.

3 Proposed Approach

As previously mentioned, our proposed approach extends the original work in [2]. In the following self-contained Section 3.1, we provide an overview of the steps involved in this disambiguation approach. In Section 3.2, we present relevant WordNet concepts that are used in our work. Finally in Section 3.3 we present our proposed approach of exploiting WordNet synsets in the disambiguation process.

3.1 WIN Algorithm

In [2], the authors propose an incremental process to disambiguate named entities. The main idea of this approach is to use the identifiers of already disambiguated entities to help in the disambiguating of later entities.

An input text document is first passed through a preprocessing phase for entity recog-

nition and entity coreference resolution (identifying each group of mentions that actually refer to the same entity). The second phase is disambiguation where the system is required to link each mention recognized in the text to its corresponding entity in Wikipedia. The disambiguation phase consists of heuristics and statistics stages; we focus on the stage of statistics where the ranking of possible Wikipedia candidate articles is performed.

In the statistics stage, a bag-of-words representation of text is employed to compute cosine similarity between two types of vector: vector representing a mention in the text and vector representing a candidate Wikipedia article. The higher the cosine similarity is, the higher the rank that candidate receives. The candidate with highest rank is chosen as the "correct" mapping for this mention into Wikipedia.

Each mention in the text can appear in a certain context (in other words, a specific position in the text). The bag of words representing a mention exploits the following features extracted from the text:

1. Local words surrounding the mention: all tokens in a window of 55 tokens surrounding the examined mention.
2. Local words surrounding coreference mentions.
3. All mentions in the document.
4. Identifiers of already disambiguated mentions: in the case of Wikipedia KB, such an identifier is the title of an article page.

The bag of words representing a Wikipedia article of an entity includes the following features:

1. Title of the entity page.
2. Title of redirect pages.
3. Category labels: a list of categories that this entity belongs to.
4. Outgoing links: titles of the entity pages that these links point to are used.

Extracted features for a mention in a text or an article in Wikipedia are then put into a bag-of-words after a normalizing process.

3.2 WordNet lexical database

WordNet [6, 7] is a large lexical database for English organized in synonym sets (synsets). As of version 2.1, WordNet contains about 150,000 words placed in over 115,000 synsets. Nouns, verbs, adjectives and adverbs are grouped into sets of synonyms – words that share the same concept and can be used interchangeably in a certain context.

Each word in a text may be mapped to a word form in WordNet if existing in the ontology. All word forms or synonyms in the same synset share the same concept and can be represented by a WordNet sense. For instance, in the 4th sense of the noun "*good*", the two word forms "*good*" and "*commodity*" express the identical meaning of articles of commerce. Thus these words constitute part of a synset with its corresponding sense identifier (sense ID).

Figure 1 shows a comparison of how Wikipedia (on the left side) and WordNet (on the right side) visually differentiate among possible entities or word senses for the mention "*plant*". In the Wikipedia's disambiguation page there is a list of links that point to articles representing different types of plant, such as "*plant*" as a living organism (a tree), or "*plant*" as a chemical or physical factory, etc… On the other hand, for all possible POS (part-of-speech), WordNet provides a synset and the corresponding definition for each sense (meaning) of "*plant*". In this case, there are four senses for the noun and six senses for the verb "*plant*".

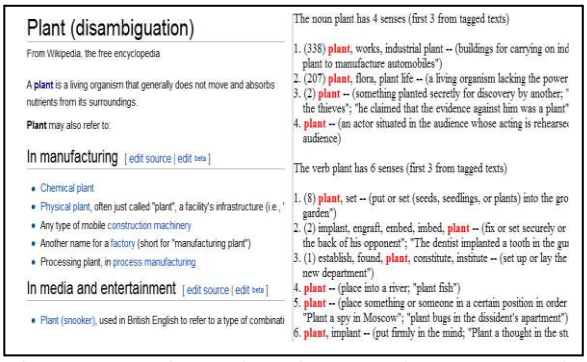


Figure 1. Information for mention "*plant*" in Wikipedia versus WordNet.

We summarize below special situations that could be met in employing WordNet and the corresponding rules to be applied. Let the word to be examined be *w*, and we wish to find its synset.

- If *w* is not found in WordNet, nothing is returned. In other words, *w* is skipped.
- If *w* is found in WordNet, and there is a number of possible senses for it (*w* is polysemous): the sense with the largest score produced from the word sense

disambiguation (WSD) process is chosen. The synset corresponding to this sense is returned.

- If *w* is found in WordNet and there is only a single sense for it (*w* is monosemous): there is no WSD processing needed and the synset corresponding to this sense is returned.

3.3 Synset expansion approach

Our proposed approach concentrates on using WordNet to modify the construction of the bag of words in the statistics stage of the WIN algorithm.

The intuition behind this approach is that due to linguistic matters, authors of various documents might choose to use different versions – or synonyms – of a word sense in the process of creating the documents. Consequently, the approach of using only observed surface forms in a body of text can lead to mismatches when we perform similarity computation. For instance, surface forms "*airplane*" and "*aeroplane*" both indicate the flying vehicle and should indeed be considered the same in the matching process.

From this observation we propose an approach to enrich the provided text features for the mention. Consider the text features representing a mention (previously described in Section 3.1). Instead of only passing that text through a normalizing process (in particular, removing stop words, punctuation, special characters and stemming) to build a bag, we construct a new version of the bag using specific WordNet features. The type of transformation examined in our approach is expansion.

In this expansion approach, for each word in the bag of words, its set of synonyms (its synset) needs to be determined and then added as an expansion to the original bag. In particular, by using synset expansion, multiple surface forms that refer to the same sense can have a higher chance to match in the ranking process.

Due to the following two reasons, only the bag of words for the mention is considered in the expanding process. The first reason is when a word is expanded into its possible synset of equivalent surface forms, it is sufficient to facilitate matching with those forms in the other side, namely, Wikipedia text, if they actually exist there. Consequently, it is not necessary for the expansion of both the bag of words representing the mention and the bag of words representing Wikipedia article.

The second reason is the processing required to extend on the Wikipedia side. In fact, we have to consider the body of Wikipedia article in order to obtain proper context for the outgoing links. Although this processing can be performed offline for Wikipedia articles, in our experiments we observe that this approach offers no positive result to the system precision. We therefore only concentrate on the approach of expanding the bags of words for mentions in the text.

Our approach consists of two methods to explore on the basic idea of expanding the bag of words for a mention in the text. The first method performs synset expansion on all words in the bag and the second method performs synset expansion only on single-sense words.

In subsequent descriptions of the two proposed methods, assume we are examining a mention *m* and its corresponding list of *k* Wikipedia candidate articles T_i, $i = 1..k$. The bag of words representing text features for *m* is $bw(m)$ and the bags of words representing Wikipedia features for Wikipedia articles are $bw(T_i)$, $i = 1..k$. Also note that there is a threshold defined to determine whether a similarity value is sufficient for the examined candidate to be the disambiguation result for *m*. The rules that handle special cases of determining word senses, as stated in Section 3.2, are applied. For example, a word is skipped when its sense cannot be found in WordNet.

In the first method, the following steps are performed:

1. For each word in the text features representing *m*, find its WordNet sense ID. Find the corresponding synset for this sense ID and replace the original word by this synset.
2. Construct the expanded version of the bag of words $bw(m)$ using the synsets found above.
3. Perform cosine similarity computations between $bw(m)$ and each of $bw(T_i)$, $i = 1..k$ to find the candidate with highest similarity to *m*.
   If this maximum is larger than the threshold, return this candidate as the mapping result for *m*.

For instance, take a simple case where the text contains "…*he lifted the band*…". The bag of words constructed is "*.. lift band* …". Then the

WSD algorithm determines that in this context, "*lift*" points to the 1st sense of the verb "*lift*" and its corresponding synset consists of ("*raise", "lift", "elevate"*). Also "*band*" points to the 4th sense of the noun "*band*" and its corresponding synset is ("*band", "banding", "stripe"*). In this case the expanded version of the bag of words is "…*lift raise elevate band banding stripe*…".

The second method presents a change made to the first step in the first method. This time we only find the synsets for words that have exactly one sense in WordNet. As previously discussed in Section 3.2, because such words have only a single sense, no WSD processing is required to determine its sense ID in order to retrieve the correct synset. As no WSD is involved, noises introduced from the WSD process are not included in the text features. The synsets for these single sense words are then used to expand the bag of words for the mention in the text.

After we have the expanded version of the bag of words, the rest of the process is similar to the original cosine similarity computation. The Wikipedia candidate with Wikipedia features that achieves highest similarity score to text features of the mention is chosen as the linking result.

We employ the tf*idf scheme to assign weights to terms in a bag of words. For the case of the mention in the text, the weighting scheme is altered to accommodate the adding of synonyms. In particular, the length of the computed vector representing the bag of words should not be affected by the added synonyms. This is because these synonyms share the same meaning with their original words in the text and therefore should not contribute to the length.

4 Experiment Results

All experiments are performed on the D_2 dataset of text documents in [2], which contain about 2,000 named entities. Obtained data are from various sources, especially scientific websites or news websites. The dataset was also employed by [2] and thus is used in order to evaluate our performance over the authors' work.

The input to the disambiguation process is a text with already recognized named entities and the output is a corresponding Wikipedia article for each named entity or NIL if there is no such article.

The WSD program in use is UKB, which utilizes a graph-based word sense disambiguation algorithm [12]. UKB applies the so-called Personalized PageRank on a Lexical Knowledge Base (LKB) to rank the vertices of the LKB and thus perform disambiguation. This algorithm requires as input a POS-tagged version of the text in order to find the senses of words. In this paper, we use the Stanford POS-tagger provided from the Stanford NLP group, and explained in [14].

The measure in which the accuracy, or performance, of the system is presented is the MAA (Micro Averaged Accuracy) measure. Let T_{all} be the number of gold mappings on a dataset (the number of correct mappings that are done manually), T_I be the number of incorrect mappings and T_C be the number of correct mappings that are done by a disambiguation system. Note that $T_{\text{all}} = T_{\text{C}} + T_{\text{I}}$. Definition of the measure is provided in the following formula:

$$MAA = \frac{T_C}{T_C + T_I} = \frac{T_C}{T_{all}}$$

Table 1 below summarizes the MAA results of the methods experimented. E1 and E2 indicates the first method (synset expansion of all words) and second method (synset expansion of only single-sense words), respectively. The ALL, PER, LOC and ORG columns contain MAA result for all named entities and for each of the three categories of named entities: person, location and organization.

Table 1. MAA results for expansion methods

Method	ALL	PER	LOC	ORG
WIN	90.1	92.4	86.8	91.4
E_1	91.6	93.3	88.1	93.2
E_2	92.5	93.9	89.3	94.8

The results show that our two proposed methods utilizing the expansion approach are able to improve the precision of the named entity disambiguation approach from [2]. In the second method, a more special criterion is employed, where only words with a single sense in WordNet are considered in the expansion process. This case may seem to be limited, but it is motivated by the fact that synonymous words added as expansion into the bag of words in this manner have the lowest change of being incorrect. In other words, only relevant information may be added and thus can serve as a more appropriate source

of enriching data.

In the first method, all words in the text features are considered in the expansion process, which may lead to irrelevant information because of the WSD system's inability to correctly identify word senses. These wrongly identified word senses are then used to expand incorrect synonyms into the bag of words and the ranking of Wikipedia candidate articles is also affected. Taking these results and analyses into consideration, a WSD algorithm that can offer higher accuracy can help to further improve the performance of the disambiguation process.

We examine the following example case of disambiguating the mention "*AP*" from an excerpt of a test document (shown in Figure 2). In this case, the original WIN algorithm maps the mention to NIL due to an insufficient amount of text features to use in similarity computation.

> GENEVA, Switzerland (**AP**) - The European Union said Monday it would be willing to slash farm tariffs by 60 percent as part of a new global trade pact, the deepest cut it has ever offered.
> EU Trade Commissioner Peter Mandelson told reporters at the World Trade Organization (WTO) that the offer was meant to kick-start a week of crunch-time talks on a new global commerce pact.

Figure 2. Excerpt from a news document containing the mention "*AP*"

Our first method uses synset expansion of all words in the text features for the mention. Although the text features are enriched, the mapping result is incorrectly determined as "*Andhra Pradesh*", because irrelevant synsets retrieved and added affect the similarity computation and consequently lead to the wrong result. Meanwhile, our second method relies on synset expansion of only single-sense words (e.g. "*reporter*" expanded into "*reporter*", "*newsman*", "*newsperson*") and helps to increase the similarity with the text features representing the correct Wikipedia article, namely "*Associated Press*".

5 Conclusion

We have explored and experimented different ways to exploit WordNet features for named entity disambiguation. The conducted experiments show that our proposed expansion method, utilizing synsets to enrich the text features for the mention, improves the performance of the system in terms of the MAA measure. The strategy that uses only synsets of monosemous words achieves higher performance by avoiding adding irrelevant information to the bag of words model.

In addition to named entities, we are currently investigating methods to use WordNet features for efficient abstract entity disambiguation. The task of disambiguating abstract entities, however, is challenging due to the complex and overlapping nature of the corresponding set of Wikipedia articles.

References

[1] Cucerzan, S.: *Large-Scale Named Entity Disambiguation Based on Wikipedia Data*. In: Proc. of Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 708-716, 2007.

[2] Nguyen, H.T., Cao, T.H.: *Named Entity Disambiguation: A Hybrid Approach*. International Journal of Computational Intelligence Systems, 5 (6), 1052-1067, 2012.

[3] Nguyen, H.T., Cao, T.H, Nguyen, T.T., Vo, L.T.T.: *Heuristics- and Statistics-based Wikification*. In: Proc. of the Pacific Rim International Conference on Artificial Intelligence, pp. 679-682, 2012.

[4] Milne, D., Witten, I.H.: *Learning to Link with Wikipedia*. In: Proc. of ACM Conference on Information and Knowledge Management, NY: ACM Press, pp. 509-518, 2008.

[5] Mihalcea, R., Csomai, A.: *Wikify!: Linking Documents to Encyclopedic Knowledge*. In: Proc. of the 16th ACM Conference on Information and Knowledge Management, pp. 233-242, 2007.

[6] Fellaum, C.: WordNet: *An Electronic Lexical Database*. MIT Press, Cambridge MA, 1998.

[7] Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.J.: *WordNet: An On-line Lexical Database*. International Journal of Lexicography, Vol.3, 235-244, 1990.

[8] Bunescu, R., Paşca, M.: *Using Encyclopedic Knowledge for Named Entity Disambiguation*. In: Proc. of the 11th Conference of the European Chapter of the Association for Computational Linguistics, pp. 9-16, 2006.

[9] Ji, H., Grishman, R., Dang, H. T.: *An Overview of the TAC2011 Knowledge Base Population Track*. In: Proc. of Text Analysis Conference, 2011.

[10] Ngo, V.M., Cao, T.H., Le, T.M.V.: *Combining Named Entities with WordNet and Using Query-Oriented Spreading Activation for Semantic Text Search*. In Proc. of the 8th IEEE International Conference RIVF, pp. 13-18, 2010.

[11] Medelyan, O., Witten, I.H., Milne, D.: *Topic Indexing with Wikipedia*. In: Proc. of the AAAI Workshop on Wikipedia and Artificial Intelligence, pp. 19-24, 2008.

[12] Eneko, A., Aitor, S.: *Personalizing PageRank for Word Sense Disambiguation*. In: Proc. of the 12th Conference of the European chapter of the Association for Computational Linguistics, pp. 33-41, 2009.

[13] Dredze, M., McNamee, P., Rao, D., Gerber, A., Finin, T.: *Entity Disambiguation for Knowledge Base Population*. In: Proc. of COLING, pp. 277-285, 2010.

[14] Toutanova, K., Klein, D., Manning, C.D., Singer Y.: *Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network*. In: Proc. of HLT-NAACL, pp. 252-259, 2003.

[15] X. Han, L. Sun, J. Zhao: *Collective Entity Linking in Web Text: A Graph-Based Method*. In Proc. of the 34th Annual ACM SIGIR Conference, pp. 765-774, 2011.

Evaluation of Web Pages Using Wikipedia

Shoki Hirata **Hidekazu Yanagimoto** **Michifumi Yoshioka**
School of Engineering, Osaka Prefecture University
Sakai, Osaka, Japan 599-8531
hirata@sig.cs.osakafu-u.ac.jp, {hidekazu,yoshioka}@cs.osakafu-u.ac.jp

Abstract

Web search engine is used widely by users that have various kinds of information needs. Hence, it is difficult to satisfy their needs with single ranking. Therefore, we have to rerank results according to the needs of each user. In this study we develop a reranking algorithm to satisfy users that want to know meaning of a query. To achieve this aim we estimate terms that are related to the query using Wikipedia. We evaluate web pages using related terms and sort web pages according to their evaluation scores. In a conventional method only headwords of Wikipedia articles are used as related terms. However, it is insufficient to satisfy user's needs. In our proposed method we use not only headwords but also terms extracted from bodies of Wikipedia articles as related terms. We confirmed the proposed method could achieve higher accuracy than the conventional method by experiments. In the experiment we made a manual ranking as labeled data and confirmed the proposed method could get stronger positive correlation with the manual ranking than the conventional method.

Keywords: Natural Language Processing, Ranking Algorithm, Wikipedia Mining

1 Introduction

Although web search engine is used under various needs, the search result is only one kind of ranking of web pages, and it is difficult to satisfy every user's needs. Thus, we should modify the ranking by the needs of each user to improve user's satisfaction. In this study we assume a situation that a user looks up an unknown word, and we aim to find web pages that have detailed explanations of his/her query.

Web pages that do not have explanation of a query come to high rank of the search result. For example, when we search "Natural Language Processing" with a search engine, we find some web pages on a book about natural language processing or web pages of laboratory making a study of natural language processing. There are few explanations about "Natural Language Processing", and users cannot know the detail of the query by viewing these web pages. Hence, it is necessary to sort the search results so that users can efficiently find web pages that serve our purpose. However, it is difficult to accomplish this aim by conventional approaches, like link-based analysis or content-based methods that measure how the query matches web pages. We estimate terms related to the query using Wikipedia and evaluate web pages using these related terms. Finally, we sort web pages according to their evaluation scores.

There is a conventional method[1] related to our study. This method evaluates web pages with respect to the coverage: how many topics related to a query a web page covers. First it analyzes link structure in Wikipedia articles and detects domain of the query. Then this method extracts terms from the detected domain. These terms are related to the topics about the query. Finally, it calculates the coverage by analyzing how many related terms are included in the web page.

Even if a web page calculated by the conventional method gets high score, the web page sometimes hardly explains the query. For example, a web page that consists of a table of contents of a book about the query includes many related terms and then gets a high score. However, a user can hardly understand the query when the user reads only the table of contents. Therefore, it is necessary to determine how much explanation about the query is included in web pages.

The conventional method uses only headwords of Wikipedia articles as related terms. However, it is not enough to find web pages that explain a query. Bodies of Wikipedia articles are explanations about their headwords. In our proposed method we extract descriptive expressions about the query from a body of an article. Then

we revise a score of the conventional method so that the web pages that contain a lot of descriptive expressions can get a high score.

Finally, we confirmed that the proposed method could achieve higher accuracy than the conventional method by experiments.

2 Conventional Method

2.1 Detection of Domain

First of all, a conventional method detects a domain of a query. It finds out categories of Wikipedia related to the query from link structure of articles. The domain consists of these categories. First the conventional method finds a Wikipedia article that headword is the query. The categories including the article of the query q is defined as $C_{direct}(q)$. In addition, there are some categories related to the query indirectly. These categories are called $C_{indirect}(q)$ in the conventional method, and it acquires them by the following methods. The scores of all categories of Wikipedia is calculated from the link structure of articles:

$$Score(c) = \frac{CF(c)}{Size(c)} \tag{1}$$

$CF(c)$ is the number of articles that is included in the category c and have a link to the article of the query. $Size(c)$ is the number of articles included in the category c. If $Score(c)$ is larger than a threshold α and $Size(c)$ is larger than a threshold β, the category c is included in $C_{indirect}(q)$. α is set as 0.5 and β is set as 5.

$Domain(q)$, the domain about a query q, is determined below:

$$Domain(q) = C_{direct}(q) \cup C_{indirect}(q) \tag{2}$$

$Domain(q)$ consists of a set of categories related to the query. Table 1 shows examples of $C_{direct}(q)$ and $C_{indirect}(q)$.

2.2 Related Terms Extraction and Web Pages Evaluation

Next, the headwords of Wikipedia articles in $Domain(q)$ are acquired as related terms of a query q. Nakatani et al[1] assume that the more frequently these related terms are used in $Domain(q)$, the more typical in $Domain(q)$ they are. Then $Typicality(t, q)$, typicality of the related term t of a query q, is calculated as follows:

$$Typicality(t, q) = \frac{LF(t, D_q)}{|D_q|} \tag{3}$$

$|D_q|$ is the number of articles included in $Domain(q)$. $LF(t, D_q)$ is the number of articles that have link to t and are included in $Domain(q)$.

If a web page includes a lot of typical terms in $Domain(q)$, the web page can cover topics about a query widely. From this idea, $Coverage(p, q)$, a score of a web page p about a query q, is calculated as follows:

$$Coverage(p, q) = \sum_t C(t, p) \cdot Typicality(t, q) \tag{4}$$

t is a related term of a query, and $C(t, p)$ is a value that takes 0 or 1 depending on whether a web page p contains t or not. $C(t, p)$is set as a binary value instead of the term frequency of t since a web page covers a topic about t when t appears in the web page at least once.

3 Proposed Method

The conventional method uses only headwords of Wikipedia articles as related terms. However, the headwords are specialized terms. Therefore, it is difficult to judge whether a web page has explanation about a query or not by only headwords. We consider that bodies of Wikipedia articles contain descriptive expressions. Thus, we use not only headwords but also terms extracted from bodies of articles as related terms in our proposed method. We think web pages containing many descriptive expressions should get a high score. We revise the evaluation score of the conventional method.

3.1 Explanatory Expressions

Bodies of Wikipedia articles are explanations about their headwords. Hence, we get descriptive expressions from the bodies of articles.

First, we detect $Domain(q)$ of a query q in the same way as the conventional method, and then we get all bodies of articles in $Domain(q)$. These bodies contain expressions used for explanation about a query q.

We extract terms that appear frequently in the bodies. These terms are considered descriptive expressions. Hence, we regard them as related

Table 1. Examples of $C_{direct}(q)$ and $C_{indirect}(q)$

Query:q	$C_{direct}(q)$	$C_{indirect}(q)$
Morphological analysis	Algorithm Morpheme Morphology Natural language processing	Full-text search
Natural language processing	Natural language processing Linguistics	Natural language processing Individual linguistics
Pattern recognition	Artificial intelligence Classification algorithm Psychology Information processing Retrieval	(None)

terms. Using Japanese morphological analysis tool MeCab[2], we extract nouns, nouns derived from verbs and adjectives. Some unsuitable terms exist in them: pronouns, suffix words, and non-independence nouns. We consider these terms are not likely to be descriptive expressions, and we exclude these terms.

t^+ is the terms extracted by the above method. To find frequently used terms we calculate each score of t^+ as follows:

$$Score(t^+, q) = \frac{TF(t^+, D_q)}{|D_q|} \tag{5}$$

$|D_q|$ is the number of articles included in $Domain(q)$, and $TF(t, D_q)$ is the frequency of t^+. If the score is larger than a threshold γ, we adopt t. We empirically set the γ as 0.2.

3.2 Revise Score

We consider web pages containing many extracted terms has explanation of a query. We revise $Coverage(p, q)$, the score of the conventional method.

$$Score(p, q) = Coverage(p, q) \cdot \sum_{t^+} C(t^+, p) \tag{6}$$

$C(t^+, p)$ is a value that takes 0 or 1 depending on whether t^+ appears in a web page p or not. If a sentence including a related term appears very frequently in a web page, $C(t^+, p)$ increases enormously. To avoid this situation we count only once although a term appears several times in a web page. The more terms a web page contains, the larger $\sum_{t^+} C(t^+, p)$ is.

4 Experiment

We searched five kinds of queries with Google search. We applied the proposed method and the conventional method to top 10 web pages in the search results and sorted web pages in score order. We also sorted the same 10 web pages manually and use them as labeled data. We verified the accuracy of the proposed method by caluclating Spearman's rank correlation coefficient. In the experiment, we used the data of Wikipedia in December 2012.

4.1 Descriptive Expressions

We show a part of terms chosen as the related terms of a query in Table 2 with our proposed method. The scores are calculated by equation (5) as showed in the third chapter. In this study we adopted the related terms with a score more than 0.2. It is thought that almost all the terms used for explanation can be acquired in the domain of each query.

4.2 Accuracy of the Evaluating Web Pages

We calculated Spearman's rank correlation coefficient between the manual ranking and the original ranking, the manual ranking and the ranking of the conventional method, and the manual ranking and the ranking of the proposed method. Spearman's rank correlation coefficient is a non-parametric measure to measure the similarity between two rankings.

Table 3 shows the Spearman's rank correlation coefficient between the manual ranking and the other 3 rankings for 5 queries. The proposed method got stronger positive correlation with the manual ranking than the conventional method,

Table 2. Descriptive Expressions and Scores

Query	Expressions	Score	Query	Expressions	Score
Tf-idf	algorithm	5.81	Natural language processing	language	9.45
	case	3.19		meaning	1.65
	palculation	2.14		study	1.55
	problem	2.02		grammar	1.55
	data	1.94		character	1.30
	analysis	1.91		case	1.16
	…	…		…	…

Table 3. Spearman's rank correlation coefficient

Query	Google	The conventional method	The proposed method
Morphological analysis	0.40606	0.79394	0.84242
Tf-idf	0.00606	0.64848	0.78182
Natural language processing	0.04242	0.28485	0.28485
Pattern recognition	0.26061	0.44242	0.52727
Neural network	-0.22424	0.15152	0.47879
Average	0.09818	0.46242	0.58303

and we found that the proposed method got more descriptive web pages. In addition, the ranking of Google search got small correlation with the manual ranking, and it is thought that the Google ranking is not suitable ranking for the purpose of this study.

5 Conclusion

In this paper we purposed a method to sort search results to find web pages that explain a query. Moreover, to achieve it we proposed a method to extract expressions used for explanation in the domain of a query from bodies of Wikipedia articles.

A problem is that this method does not cope with a query with ambiguity. When a query has ambiguity, $Domain(q)$ becomes vague and cannot acquire appropriate related terms. We will be able to overcome this problem by developing a system that user appoints an appropriate $Domain(q)$ by showing the candidates of $Domain(q)$.

This method evaluates only one aspect of web pages. We will develop a system that can evaluate other aspects of web pages in the future.

References

[1] M.Nakatani, A.Jatowt, H.Ohshima, and K.Tanaka. "Quality Evaluation of Search Results by Typicality and Speciality of Terms Extracted from Wikipedia", In Proceedings of the 14th International Conference on Database Systems for Advanced Applications (DASFAA 2009), pages 570-584, 2009.

[2] T.Kudo, K.Yamamoto, Y.Matsumoto. "Applying Conditional Random Fields to Japanese Morphological Analysis", Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, pages 230-237, 2004.

Evaluative Phrase Extraction from Customer Review

Satoshi Sakemi **Hidekazu Yanagimoto** **Michifumi Yoshioka**
School of Engineering, Osaka Prefecture University
Sakai, Osaka, Japan 599-8531
{sakemi@sig.cs, hidekazu@cs, yoshioka@cs}.osakafu-u.ac.jp

Abstract

When we want to buy products on an EC site, we can get some information from merchandise pages and customer reviews. In merchandise pages, there is information provided by producers. On the other hand, we can get subjective information wrote by a customer from customer reviews. Subjective information contains relative comparison or subjective evaluation made by customers, and it is different depending on products. Therefore, we cannot decide evaluative phrases previously. In this study, the main aim is extracting evaluative phrases from customer reviews automatically. We use the dependency analysis to achieve the aim. Furthermore, we group similar evaluative phrases by using similarities of dependency. Then we judge the similarity using a graph analysis method. Speaking concretely, we represent relation between noun and adjective as a bipartite graph. The word similarity is defined using the bipartite graph similarity. Using this method, we confirmed that evaluative phrases are classified into appropriate categories according to their meanings.

Keywords: Natural Language Processing, Text Mining

1 Instruction

In these days, a market scale of electronic commerce has grown rapidly. Accordingly, we can buy various products on the Internet. When we use an EC site, we can get information that is necessary for decision making from merchandise page and customer reviews. In merchandise page, there are performance sheets and specification sheets provided by producers. On the other hands, a customer review includes impression or a valuation of products reported by customers who has purchased them and is introduced in many EC sites. The information we can get from customer reviews bases on subjectivity of a customer so we call this subjective information. Subjective information is important factor for making a decision but products have various aspects and the aspects regarded as important are vary from genre to genre. Hence, it is difficult to decide evaluative phrases previously.

In this study the main aim is extracting evaluative phrases from customer reviews automatically. To be concrete, a noun modified by an adjective, which tend to contain evaluation, is considered as a candidate of evaluative phrase. Evaluative phrases are selected from the candidates based on occurrence frequency. Furthermore, we construct bipartite graph to represent relation between noun and adjective and calculate similarities according to relation strength between nouns and adjectives. Then we attempt to discover similar evaluative phrases by using the similarities.

2 Related Work

Sentiment analysis, which is the task of identifying polarity of natural language text is related to this study. Tony Mullen et al. proposed an approach to this task by using support vector machines[1]. Theresa Wilson et al. reported an approach to phrase-level sentiment analysis[2]. These works only estimate polarity of the whole text or phrases in the text; it is not considered what aspect of a product authors evaluate. In this study, we extract an evaluated aspect of a product as a word. Namrata Godbole et al. reported another approach estimating an entity polarity[3]. In their paper, a target word is fixed, and its polarity is estimated. Hence, it is difficult to categorize words according to their usage in large corpus.

3 Evaluative Phrases Extraction

Adjectives appearing in customer reviews express features of products and subjective evaluations while nouns modified by them express aspects of the products. Therefore we con-

sider such noun as an evaluative phrase. We used Japanese dependency structure analyzer CaboCha to analyze dependency between adjectives and nouns. Basic method to extract evaluative phrases is simple; if there is a dependency between the chunk that contains an adjective and the chunk that contains a noun, the noun is a candidate of an evaluative phrase. Moreover, to improve accuracy of extraction, we used additional rules as follows:

- Deletion
  Pronouns and non-independent nouns are excluded because such words cannot make sense alone. The nouns derived from verbs or adverbs are also excluded because they have function to convey action or emphasis and they do not tend to show the aspect of products.
- Chunking
  The nouns that have function like a suffixes are combined with the verb appeared just before.

The candidates are sorted in decreasing order of their occurrence frequency. If occurrence frequency is higher than threshold, we consider it evaluative phrase. We assume that the higher occurrence frequency is, the more important the evaluative phrase is.

4 Discovering similar evaluative phrases

We attempt to classify evaluative phrases according to similarity of relation between evaluative phrases and adjectives. Dependency between evaluative phrases and adjectives is represented as a bipartite graph which weights are proportional to their occurrence frequency. Similarity is calculated by using this graph. We use Jaccard index and Jensen-Shannon divergence as similarity measures. Let $N_1, ..., N_i$, $A_1, ..., Aj$ and $w_{a,b}$ be nominal nodes, adjectival nodes and weight that edges have respectively.

4.1 Jaccard index

The Jaccard index represent similarity of two sets. Let S_a and S_b be sets, Jaccard index $J(S_a, S_b)$ is expressed in the following form.

$$J(S_a, S_b) = \frac{|S_a \cap S_b|}{|S_a \cup S_b|} \tag{1}$$

In this study, we define S_x as follows and calculate Jaccard index for all combination of evaluative phrases.

$$S_x = \{A_y | w_{x,y} \geq 1\} \tag{2}$$

As two evaluative phrases share many common adjectives, the evaluative phrases are highly related.

4.2 Jensen-Shannon divergence

Jensen-Shannon divergence is a method of measuring similarity between two probability distributions. Jensen-Shannon divergence of two distributions P_a and P_b is expressed in the following form.

$$\begin{aligned} D_{JS}(P_a||P_b) &= \frac{1}{2}\sum_x P_a(x)\log\frac{P_a(x)}{M(x)} \\ &+\frac{1}{2}\sum_x P_b(x)\log\frac{P_b(x)}{M(x)} \\ M(x) &= \frac{1}{2}(P_a(x) + P_b(x)) \end{aligned} \tag{3}$$

$P_a(x)$ and $P_b(x)$ are the probabilities that x is observed on each distribution. Let $P_1, ..., P_i$ be distributions of each evaluative phrases. Jensen-Shannon divergence is calculated for all combinations of the distributions. The probability distribution of N_y is defined as follows.

$$P_y(x) = \frac{w_{y,x}}{\sum_{k=1}^{j} w_{y,k}} \tag{4}$$

To avoid zero frequency problem, we use Laplace smoothing.

$$P_y(x) = \frac{w_{y,x} + 1}{\sum_{k=1}^{j} w_{y,k} + j} \tag{5}$$

When two evaluative phrases generate similar distributions of adjectives, the evaluative phrases are highly related.

5 Experiments

5.1 Review Data

We use 16.6 million customer reviews in Rakuten Ichiba. Rakuten Ichiba has hierarchical genre structure. When we use customer reviews that belong to a genre, we include the reviews belonging to lower layers of the genre. For example, when we get the reviews that belong to ”Computer / Peripherals”, we also take the reviews that belong to ”Computer”, ”Peripherals”, ”Display”, ”Laptop”, ”Mainframe” and so on.

5.2 Procedure for Experiment

Procedure for experiments is in the following.

1. Selecting the merchandise genre and gathering all customer reviews.

2. Extracting candidates of evaluative phrase from the reviews.

3. Sorting the candidates according to their occurrence frequency.

4. Excluding the candidates which frequency is smaller than a threshold σ. Remaining ones are regarded as evaluative phrases.

5. Calculating similarity for all combinations of the evaluative phrases and clustering evaluative phrases.

6 Results

6.1 Extracting Evaluative Phrases

We show top 20 evaluative phrases from two genres: ”Display” and ”Men’s suits”. An evaluative phrases adding the number, for example ”price(2)”, means synonym in Japanese. These tables show that evaluative phrases vary from a genre to a genre. Although ”Men’s suits” and ”Men’s jeans” belong to the same upper layer ”Men’s fashion / shoes”, there are different evaluative phrases like ”tailoring” in ”Men’s suits” and ”damage” in ”Men’s jeans” in each genre. Besides, generality of the evaluative phrases change by the layer; evaluative phrases extracted from upper layer tend to be more general than those from a lower layer.

6.2 Discovering Similar Evaluative Phrases

6.2.1 Difference Between Jaccard Index and Jensen-Shannon Divergence

The evaluative phrases that are similar to ”monitor” with $\sigma = 20$ and ”price(1)” with $\sigma = 10$ are described in Table3 and Table4. From these tables, it is found that the synonyms and the related words tend to have high degree of similarity; for ”monitor”, ”display” and ”picture quality” have high degree of similarity, and for ”price(1)”, ”price(2)” and ”postage” have high degree of similarity. Although we calculate similarity using Jaccard index and Jensen-Shannon divergence based on similarity of dependency, similar evaluative phrases are not always lined up in the order of the similarity of the semantic content.

When we calculate Jaccard index, we do not care about occurrence frequency of dependency. On the other hand, we use occurrence frequency to construct distribution when we calculate Jansen-Shannon divergence. Two approaches are quite difference in this point. In general, Jensen-Shannon divergence is suitable for the evaluative phrase that is high frequency, and Jaccard index is suitable for the evaluative phrases that is low frequency. If the occurrence frequency of the evaluative phrases is low, small difference of appearance of the dependency have a serious influence on the probability distribution. This causes the difference between two approaches.

6.2.2 Effectivity of Threshold

The evaluative phrases that are similar to ”price(1)” with $\sigma = 20$ and with $\sigma = 3$ are in Table5 and Table6. When threshold is too big, the words that have relation to ”price(1)” like ”postage” and ”secondhand goods” are not considered as evaluative phrases. On the other hand, when threshold is too small, new related words like ”bargain rate”, ”points” and ”impulse buying” appear, but ”postage” and ”secondhand goods”, which are quite related to ”price(1)” with $\sigma = 3$ go down. Therefor, if we want to get only important similar evaluative phrases, we should set big σ. On the other hand, if we want to get as many similar evaluative phrases as possible, we should set small σ.

6.3 Applying the Evaluative Phrases

To confirm how the evaluative phrases appear in the customer reviews, we extract the pairs of evaluative phrase and adjective that contain evaluative phrases from the customer reviews for a product. We show a part of results we apply the evaluative phrases to the customer reviews of a product that belongs to genre ”Display”. The majority of the results are understandable, but some pairs of evaluative phrase and adjective that contain the evaluative phrase like ”feeling” and ”impression” are not understandable alone. These words do not express aspects of products. However occurrence frequencies tend to be high because these words are often modified by frequent adjectives like ”good”.

Table 1. Evaluative phrases in genre ”Display”

occurrence frequency	evaluative phrase	occurrence frequency	evaluative phrase
132	picture(1)	35	dot
122	merchandise(1)	32	PC
87	monitor	31	condition
77	price(2)	29	picture quality
71	price(1)	28	product
66	feeling	25	picture(2)
59	problem	25	viewing angle
52	liquid crystal	22	TV
45	size	20	comuputer
37	display	20	jack

Table 2. Evaluative phrases in genre ”Men’s suits”

occurrence frequency	evaluative phrase	occurrence frequency	evaluative phrase
68	suits	12	quality
57	merchandise(1)	10	gray
38	cloth	10	pants
32	feeling	8	stripe
27	sewing(1)	8	brand
23	size	7	best
22	price(1)	7	sewing(2)
22	price(2)	7	oneself
13	merchandise(2)	6	sewing(3)
13	problem	6	postage

pair of evaluative phrases and adjective

lower than used monitor
weak feeling
small in display
size is also good
good, problem
price, inexpensive
picture is clear
picture is whitish
condition that hard to see
reasonable price
price is very
clear picture
monitor that can display by full HD
want monitor
HDMI is useful
more superior than monitor
contrast is superior
using DVI, it can be detail
bad impression

7 Conclusion

We extract evaluative phrases from customer reviews in Rakuten Ichiba by using dependency of words. Moreover, we can find similar evaluative phrases using Jaccard index and Jensen-Shannon divergence. However, there is a problem that unsuitable words such as ”feeling” and ”impression” are also extracted. To improve accuracy of extraction by excluding these words will become an issue in the future.

References

[1] Tony Mullen and Nigel Collier. Sentiment analysis using support vector machines with diverse information. In:ACL(2004)

[2] Theresa Wilson, Janyce Wiebe and Paul Hoffmann. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In:HLT(2005)

[3] Namrata Godbole, Manjunath Srinivasaiah and Steven Skiena. Large-Scale Sentiment Analysis for News and Blogs. In:ICWSM(2008)

Table 3. similar evaluative phrases with ”monitor” $\sigma = 20$

Jaccard		JSD	
degree of similarity	evaluative phrase	degree of similarity	evaluative phrase
0.315	liquid crystal	0.0526	size
0.294	size	0.0544	condition
0.279	merchandise	0.0562	product
0.260	display	0.0563	liquid crystal
0.259	price(1)	0.0619	picture quality

Table 4. similar evaluative phrases with ”price(1)” $\sigma = 10$

Jaccard		JSD	
degree of similarity	evaluative phrase	degree of similarity	evaluative phrase
0.302	liquid crystal	0.0474	price(2)
0.293	price(2)	0.0671	secondhand goods
0.268	monitor	0.0754	postage
0.263	merchandise	0.0848	liquid crystal
0.231	product	0.0932	HDMI

Table 5. similar evaluative phrases with ”price(1)” $\sigma = 20$

Jaccard		JSD	
degree of similarity	evaluative phrase	degree of similarity	evaluative phrase
0.302	liquid crystal	0.0474	price(2)
0.293	price(2)	0.0848	liquid crystal
0.268	monitor	0.104	monitor
0.263	merchandise	0.114	merchandise
0.231	product	0.128	picture quality

Table 6. similar evaluative phrases with ”price(1)” $\sigma = 3$

Jaccard		JSD	
degree of similarity	evaluative phrase	degree of similarity	evaluative phrase
0.302	liquid crystal	0.0380	low price
0.293	price(2)	0.0434	HD
0.268	monitor	0.0465	point
0.263	merchandise	0.0474	price(2)
0.231	product	0.0539	impulse buying

Using Core Phrases and Information Features for Keyphrase Extraction

Huyen K. Nguyen† **Tru H. Cao†***

†University of Technology – VNUHCM, Vietnam
*John von Neumann Institute – VNUHCM, Vietnam
nguyenkimhuyen@ymail.com, tru@cse.hcmut.edu.vn

Abstract

Keyphrases express the main contents of a document. In this paper, we focus on semantic relations between phrases. The main idea is to use initial phrases to find keyphrases through semantic relations between them. In this paper, we introduce two methods to identify initial phrases: Core Phrases and Information Features methods. The Core Phrases method extracts phrases from the titles and sentences extending information of the titles. The Information Features method uses two important features in keyphrase extraction: term frequency and first occurrence position to find phrases contributing main topics of a document. Our methods outperform SemiRank and two derivatives of KEA and KEA++ on F1 measure. In addition, we show that, in comparison to using initial keyphrases to find new ones, the new methods give better results in the case that initial keyphrases are re-ranked based on their semantic relations.

Keywords: semantic relatedness, semantic relation, semantic hyper-graph, term frequency, first occurrence position.

1 Introduction

Keyphrases are single or multiple words summarizing the main contents of a document. Keyphrases are usually manually assigned by authors or experts and applied to document management systems. However, with the explosion of information sources, manual document management becomes more and more infeasible. Consequently, automatic extraction of keyphrases is required. Keyphrases are also applied to many fields of natural language processing (NLP) such as clustering and question answering.

There are two main approaches for keyphrase extraction: supervised and unsupervised learning. Supervised methods [1,2,3,4] build statistical models from training datasets for ranking the importance of candidate keyphrases. Unsupervised methods [5,6,7] define keyphrases' features in formulas for weighting phrases. However, semantic relations between phrases have not been adequately considered in both approaches. Instead, they focus on the information that phrases contribute to a text. In this paper, we pay attention to semantic relations between phrases, in particular the semantic relatedness between them.

Our research is based on an observation that keyphrases assigned by annotators have close semantic relations to others. For example, once "*Machine learning"* is chosen as a keyphrase, phrases semantically close to it such as *"Statistical classification"* or "*Artificial intelligence"* are also selected if they appear in the document. Naturally, that makes us look for methods to find new keyphrases when we have other ones.

[8] has already investigated this idea. The work uses phrases from the title as initial keyphrases and finds others based on semantic relatedness between them. However, the number of phrases in a title is limited. In some documents' titles, there is no more than one meaningful phrase.

In this paper, we extend the number of initial keyphrases by two methods: Core Phrases and Information Features. The Core Phrases method extracts important phrases from titles and from sentences extending information of those titles. The Information Features method finds important phrases by two efficient features in keyphrase extraction: term frequency (TF) and first occurrence position (FOC) [5,7,9]. TF and FOC represent the degree of information that a phrase contributes to the contents of a document. In addition, we show that instead of using initial keyphrases to refer to new ones, an algorithm gives better results in case that we re-rank the importance of initial keyphrases.

The rest of the paper is organized as follows. In Section 2, we give an overview about the

method of [8] and present our two proposed methods to extract initial keyphrases. In Section 3, we discuss testing datasets and preprocessing. We evaluate experiment results in Section 4 and give concluding remarks in Section 5.

2 Proposed Methods

This section is separated into three parts. First, we introduce the method of [8] (henceforth referred to as SemiRank). The two next subsections present our methods for improving a set of initial keyphrases. The first method finds main phrases from a title and sentences with high relations to the title in the same document. The second method captures keyphrases based on two well-known features in keyphrase extraction.

2.1 SemiRank

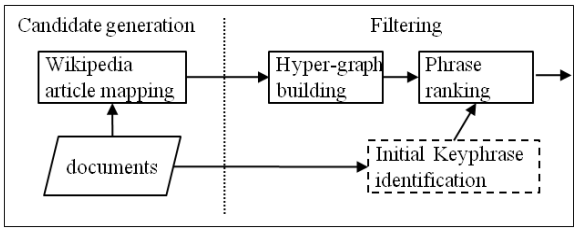


Figure1.Processing steps of SemiRank

The main idea of SemiRank is that from a small number of initial keyphrases it uses semi-supervised learning to assign the keyphrase role to other phrases in a document. The new keyphrases have close semantic relations to the initial ones.

Figure1 shows the process of SemiRank. Like other keyphrase extraction systems [1,5,7,9], it includes two stages: candidate generation and filtering. The first stage analyses the text of a document to generate a set of candidate keyphrases, then the second stage computes their weights based on pre-defined features. Phrases with the highest scores are selected as keyphrases.

In SemiRank, the first stage extracts and matches a phrase in a document to a Wikipedia article expressing its meaning. The WSD (Word Sense Disambiguation) technique from [10] is applied to resolve ambiguity in mapping.

In the first step of the second stage, for each document, a weighted hyper-graph is built as a semantic network where each vertex is a phrase assigned with an appropriate Wikipedia article about its meaning. Each edge represents the semantic similarity of the two respective phrases. There are both binary edges and *n*-ary edges in the hyper-graph. The weight of a binary edge is measured by the semantic relatedness metric in [10] based on Dice's measure to capture common links between the two corresponding Wikipedia articles. For defining *n*-ary edges, the method proposed by [11] is applied to group the vertices into a community, where each community is a hyper-edge. Phrases in a community have close relations to each other and less close to phrases in the other communities. The weight of a hyper-edge is the average weight of all binary edges whose vertices belong to the community.

In the second step, titles are used for extracting important phrases. Then, each phrase is matched to an appropriate Wikipedia article. These phrases are considered to be initial keyphrases and are input for the ranking algorithm in the last step.

In the last step, the ranking algorithm, named PhraseRank, based on the framework proposed by [12] is performed. At the initialization of the algorithm, the initial keyphrases are assigned value 1 and the other phrases are assigned value 0. PhraseRank travels recursively through all phrases in semantic network to determine their scores. The purpose of the algorithm is to recursively spread the score of a phrase to its neighbors. The value that the phrase assigns for a neighbor depends not only on its score but also on their semantic relation. If they have close semantic relatedness, which means the weight of the corresponding binary edge is high and they are in the same community, then the value assigned to the neighbor of that phrase is high.

After ranking, the *k* top ranked phrases are selected as keyphrases.

In Figure 1, the dashed line box is the modified component of SemiRank in our work. The next subsections present our methods for identifying initial keyphrases based on core phrases and information features.

2.2 Core Phrases Method

SemiRank votes phrases extracted from a title as keyphrases since titles usually express the top main contents of a document. Relations between keyphrases of a document and its title are also mentioned in [3,7]. In [3], the statistics shows

that the density of keyphrases in titles is the highest in comparison to the other sections (e.g. abstract, introduction …) of a document. In the expression used to weight keyphrases, [7] assigns the highest score to phrases that belong to a title. Those are reasons why we try to find sentences extending information in a title.

To extract such sentences, we filter each sentence in a document under two conditions:

1. They have to be in the abstract, introduction or conclusion sections of a document.
2. In our observation in scientific reports, if a writer wants to emphasize the main contents of a document, they usually express it with the following special terms: *"paper/ this work/ this study/ this article/ we/ our"*. Therefore, sentences containing these terms will be selected for the next steps.

A sentence after filtering by the above conditions is weighted based on the similarity between it and the title of the document to which it belongs. Words that appear in both the sentence and the title are counted and the weight of the sentence is the ratio of their common words to the length of the sentence:

$$W_s = \frac{|T \cap S|}{|S|}$$

where T is the set of words in the title, and S is the sets of words in the sentence.

In addition, to avoid short sentences repeating information of the title, we remove any sentence whose weight is smaller than 0.5. It means that its length must be greater than, or equal to, twice the number of common words.

For example, in the document of Wiki-20 test set [9] with the title *"Occam's Razor: The Cutting Edge for Parser Technology"*, we expect to extract the sentence "*This paper provides a demonstration of this state-of-the-art compiler-compiler technology using the programming language occam as an example*" in the abstract, because it has significant phrases "*compiler-compiler*", *"programming language"* and *"occam"*. This sentence satisfies the conditions: belonging to abstract section, containing the term *"this paper"*, and its length greater than twice the number of common words (*"occam"*, *"technology"*).

These filtered sentences and the title of a document contain core phrases, i.e., important topics, of the document. After the weights of the filtered sentences are calculated, they are sorted and the k top sentences are selected.

The selected sentences and the title are then used to extract important phrases and their appropriate Wikipedia articles. The n top phrases with the highest frequencies are considered to be initial keyphrases for the next steps.

2.3 Information Features Method

In automatic keyphrase extraction systems, attributes of manually annotated keyphrases are observed and then deployed. These features are applied to build statistical models for classifying candidate keyphrases in supervised learning methods [1,2,3,9]. Meanwhile, in unsupervised learning methods [5,6,7], these features are combined into formulas to weight phrases. In this work, we use two important and efficient features of keyphrases: TF and FOC.

TF represents the number of times a phrase appears in a document. FOC refers to where a phrase appears in the first time. FOC of a phrase is calculated as the distance in words between the position of the first occurrence of the phrase and the beginning of the document.TF and FOC reflect how important a phrase is. If a phrase provides significant information of a document, the writer tends to repeat it frequently and discuss it soon in the document. TF and FOC are normalized by the length of the document.

TF and FOC are often used as features for classifying candidate keyphrases, in particular in [5,6,7,9]. However, these systems are complex in computation and implementation [5,9], as well as require additional resources [6,7]. Therefore, we proposed a simple phrase weighting equation based on these features as follows:

$$W_p = \frac{TF}{FOC}$$

After weighting all phrases, they are sorted and the k top phrases are returned as initial keyphrases.

3 Dataset and Pre-processing

The dataset Wiki-20 [9] is used for evaluating our methods. Wiki-20 contains 20 reports in many areas of computer science such as programming, algorithms, or graph theory. Each document of Wiki-20 is annotated by fifteen independent teams of graduate and undergraduate students in computer science, using Wikipedia as a con-

trolled vocabulary. Keyphrases of each document are Wikipedia article titles matching to terms extracted from the document. On average, there are 5.7 keyphrases per document.

Since the document volume in dataset Wiki-20 is large, the process of candidate keyphrase generation creates a large number of candidates. This makes the PhraseRank algorithm take long time and many CPU resources to execute. Therefore, we reduce the number of candidates using a filter with two conditions before building a hyper-graph:

1. Keyphraseness: keyphraseness is proposed by [13] to measure the importance of a phrase. Keyphraseness value of a phrase p is computed by the proportion of the number of articles where p appears as a link, denoted by $N(p = link)$, to the total number of articles where it appears, denoted by $N(p)$:
$$Keyphraseness_p = \frac{N(p = link)}{N(p)}$$
While composing an article, if it contains an important term that is explained in details in another article, the writer will create a link from the term to that article. An article that is usually linked to each time a term having a relation to it occurs (e.g. synonym, abbreviation …) should be taken more attention than others. In this paper, only phrases with keyphraseness greater than, or equal to, 0.02 are selected.
2. Long phrases: if two or more short phrases standing next to each other matching with some Wikipedia articles and the whole long phrase covering them also corresponds to an article, then we remove all short phrases and keep only the long phrase. For example, as there are three Wikipedia articles explaining the three phrases *"regression"*, *"test"* and *"regression test"*, we remove both of the two short phrases, and keep only the long phrase *"regression test"* and its corresponding Wikipedia article for the next steps.

On average, we get 232 candidate keyphrases per document.

4 Evaluations

We compare our methods with three other keyphrase extraction ones. The first one is SemiRank that uses initial keyphrases from titles only. We re-implement that method to compare the initial keyphrases extracted from titles with the ones extracted by our methods. The other methods are two derivatives of KEA and KEA++ proposed in Maui system [9]. KEA [3] is a keyphrase extraction system based on supervised learning. A Naïve Bayes model is built from three features: TF*IDF, FOC, and keyphraseness. KEA++ [14] is an improvement of KEA, introducing two more features, namely, term length and node degree, but neglecting keyphraseness. However, KEA and KEA++ extract keyphrases from a document or controlled vocabulary, but not Wikipedia. Therefore, in Maui system, KEA and KEA ++ are enhanced with mapping phrases to Wikipedia articles in the step of candidate keyphrase generation. The features of KEA and KEA++ are combined to this new generation step in the derivatives of KEA and KEA++. Their performance that we show in this work is reported in [9].

The precision, recall and F1 measures are used to evaluate the effectiveness of the methods. The measures evaluate automatically extracted keyphrases against manually selected ones. Correct extracted keyphrases are phrases in both automatic sets and manual sets. Precision is the fraction of correct keyphrases out of automatic keyphrases:
$$precision = \frac{|AKP \cap MKP|}{|AKP|}$$
where *AKP* is a set of automatic keyphrases and *MKP* is *a* set of manual keyphrases.
Recall is the fraction of correct keyphrases out of manual keyphrases:
$$recall = \frac{|AKP \cap MKP|}{|MKP|}$$
F1 is the harmonic mean of precision and recall:
$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

After performing the ranking algorithm, the five phrases with the highest scores are selected as keyphrases. We compare these top-5 keyphrases against sets of manual keyphrases of the fifteen teams using precision, recall and F1. The overall results are their average values.

4.1 Evaluation Results

Before evaluating our methods, we want to note that in the experiments with Wiki-20 test dataset, the ranking algorithm Phrase Rank assigns initial phrases higher scores than the other candidates. It

Table 1. Performance of the Core Phrases and Information Features Methods

		Core Phrases			Information Features		
		P	R	F1	P	R	F1
1	3 initial kps	19.5	17.5	18.2	19.6	17.6	18.3
2	4 initial kps	21.1	19.0	19.7	23.7	21.3	22.1
3	5 initial kps	21.5	19.2	20.0	23.7	21.3	22.1
4	6 initial kps	22.5	20.2	21.0	23.9	21.5	22.4
5	7 initial kps	23.5	21.0	21.9	24.8	22.3	23.2
6	8 initial kps	22.7	20.3	21.2	26.2	23.3	24.4
7	9 initial kps	22.5	20.1	20.9	25.8	23.0	24

means that if we have *k* initial keyphrases before PhraseRank is performed, they will be in top-*k* of final keyphrases.

In this subsection, firstly we investigate the appropriate number of sentences in the Core Phrases method. Secondly, we show the performance of our methods with different sets of initial keyphrases. Finally, we compare them with SemiRank and the derivatives of KEA and KEA++.

To determine an adequate number of sentences in the Core Phrase method, we evaluate the performance with the top-3 sentences, and then the top 4, 5, 6, 7 and 8 sentences. For each set of phrases obtained from the top-*k* sentences and the title of the same document, we select the top-*n* ones of them with the highest frequencies as initial keyphrases where *n* is ranged from 3 to 9. The ranking algorithm is performed on each top-*n* initial keyphrases. As results, the top-4 (k = 4) sentences give the best result on F1 measure with 7 initial phrases (n = 7).

The experiments with different sets of initial keyphrases are repeated for the Information Features method. The size of the initial ones is ranged from 3 to 9.The weight of an initial keyphrase is calculated as mentioned in Section 2.3. As a result, the method gives the best performance on F1 measure with 8 initial phrases.

Table 1 shows the performance of the Core Phrases and Information Features methods. The Core Phrases method uses top-4 sentences. The size of initial ones is ranged from 3 to 9. We have three different cases based on the size of initial keyphrases and the size of final keyphrases:

1. If the number of initial phrases is smaller than the number of final keyphrases, PhraseRank generates new keyphrases excluding the initial keyphrases (rows 1 and 2).
2. If the number of initial phrases is equal to the number of final keyphrases, the PhraseRank algorithm changes nothing. It means the initial phrases are the final keyphrases (row 3).
3. If the number of initial phrases is greater than the number of final keyphrases, the algorithm re-ranks the importance of initial keyphrases based on their relations (rows 4 to 7).

The Core Phrases and Information Features methods give the best results at the 7 and 8 initial phrases respectively. In this case, the number of initial keyphrases exceeds the number of final keyphrases.

At the beginning, all initial keyphrases of a document are assigned with score 1 and the others are assigned with score 0. These values of initial phrases are spread to their neighbors, and then recursively from neighbors to neighbors using their semantic relations. While the whole structure of the network is explored, the scores of all phrases, including initial keyphrases, are modified but these initial ones are added with an extra constant value [8] to ensure that they are always in top-*k* final keyphrases as mentioned earlier in this section. Therefore, the final keyphrases are the initial keyphrases assigned with the highest scores, which have closer semantic relations to other phrases in the document than the unselected initial keyphrases.

Table 2 compares the performance of the Core Phrase and Information Features methods to SemiRank and KEA and KEA++ adapted for Wikipedia. As mentioned above, the performance figures of KEA and KEA++ adapted for Wikipedia are report in [9]. In [9], inter-consistency measure, which is equal to F1, is used. In our work, inter-consistency measure is replaced by F1 measure because it is used popularly for evaluating keyphrase extraction systems.

Since inter-consistency measure does not record precision and recall, they are missed in Table 2.

In Table 2, the Information Features method is superior to all other methods at 8 initial keyphrases. It achieves 24.4% which 4.9% higher than SemiRank on F1 measure.

Table2. Performance of different keyphrase extraction methods for Wiki-20

Approach	P	R	F1
SemiRank	20.9	18.7	19.5
KEA for Wiki	-	-	20.2
KEA++ for Wiki	-	-	22.6
Core Phrases(8 initial kps)	22.7	20.3	21.2
Infomation Features (8 initial kps)	26.2	23.3	24.4

5 Conclusion

In this work, we investigate using initial keyphrases to find keyphrases based on semantic relations. We adapt the SemiRank method with extension of a set of initial keyphrases. Instead of considering only phrases in a title to be initial keyphrases, we develop two methods for extracting the initial keyphrases: the Core Phrases and the Information Features methods. Both of them outperform SemiRank and the derivatives of KEA and KEA++. In addition, we show that re-ranking initial keyphrase brings better performance. The proposed re-ranking of initial keyphrases could be applied to other keyphrase extraction systems where semantic relations have not yet been exploited.

References

[1] Witten, I.H., Paynter, G.W., Frank, E., Gutwin, C., and Nevill-Manning, C.. *KEA: Practical Automatic Keyphrase Extraction*. In: Proc. of the 4th ACM Conference on Digital Libraries, pp. 254–255, 1999.

[2] Hulth, A.. *Combining Machine Learning and Natural Language Processing for Automatic Keyword Extraction*. Ph.D. Thesis, Stockholm University, 2004.

[3] Nguyen, D.T. and Luong, T.M.. *WINGNUS: Keyphrase Extraction Utilizing Document Logical Structure*. In Proc. of the 5th International Workshop on Semantic Evaluation, ACL, pp. 166–169, 2010.

[4] Berend, G. and Farkas, R.. SZTERGAK : *Feature Engineering for Keyphrase Extraction*. In Proc. of the 5th International Workshop on Semantic Evaluation, ACL, pp. 186–189, 2010.

[5] El-Beltagy, S.R. and Rafea, A.. *KP-Miner: Participation in SemEval-2*. In Proc. of the 5th International Workshop on Semantic Evaluation, ACL, pp. 190–193, 2010.

[6] Bordea, G. and Buitelaar, P.. *DERIUNLP: A Context Based Approach to Automatic Keyphrase*. In Proc. of the 5th International Workshop on Semantic Evaluation, ACL, pp. 146–149, 2010.

[7] You, W., Fontaine, D. and Barthès, J. P.. *An Automatic Keyphrase Extraction System for Scientific Documents*. Knowledge and Information Systems, Vol. 34, 691-724, 2012.

[8] Li, D., Li, S., Li, W., Wang, W., Qu, W.. *A Semi-Supervised Key Phrase Extraction Approach: Learning from Title Phrases through a Document Semantic Network*. In Proc. of the ACL 2010 Conference Short Papers, pp. 296–300, 2010.

[9] Medelyan, O.. *Human Competitive Automatic Topic Indexing*. Ph.D. Thesis, University of Waikato, 2009.

[10] Turdakov, D. and Velikhov, P.. *Semantic Relatedness Metric for Wikipedia Concepts Based on Link Analysis and its Application to Word Sense Disambiguation*. In Proc. of SYRCODIS 2008 Colloquium on Databases and Information Systems, 2008.

[11] Newman, M.E.J.. *Analysis of Weighted Networks*. Physical Review E, Vol. 70, 056131, 2004.

[12] Zhou, D., Huang, J. and Schölkopf, B.. *Beyond Pairwise Classification and Clustering Using Hypergraphs*. MPI Technical Report, No. 143, 2004.

[13] Mihalcea, R. and Csomai, A.. *Wikify!: Linking Documents to Encyclopedic Knowledge*. In: Proc. of the 16th ACM Conference on Information and Knowledge Management, pp. 233-242, 2007.

[14] Medelyan, O.. *Automatic Keyphrase Indexing with a Domain-Specific Thesaurus*. Master Thesis, Albert-Ludwig University Freiburg, 2005.

ACIS 2013

Asian Conference on Information Systems 2013

Poster Session

Robust Jointly Sparse Elastic Regression with Matrix Uncertainty

Seppo Karrila[†]
[†]Faculty of Science and Industrial Technology
Prince of Songkla University, Surat Thani, Thailand
seppo.karrila@gmail.com

Abstract

Straightforward and practical linear programs enable robust L1 or sup-norm regression with

- interval errors-in-variables (a.k.a. matrix uncertainty)
- underdetermined systems with an excess of features to select from
- tunable solution sparsity
- jointly sparse solutions for multiple outputs (multi-task learning), and
- tunable robustness (redundancy) against the loss of some selected features.

Keywords: Compressed Sensing, Linear Interval Systems, Linear Programming, Robust Regression, Joint Sparsity

1 Introduction

We address a significant problem of modelers who train predictive models by fitting data.

When the training cases are re-measured and the training data reproduced within measurement inaccuracy, a fitted model must have relatively small changes in its outputs. Otherwise the model does not describe the cases, just the numbers that happened to be in an instance of training data. So a useful model must be robust against reasonable perturbations also in its input variables, and this consideration takes errors-in-variables into account on model identification. *Significant input inaccuracy must affect model identification.* Still, often it is ignored because of lack of tools. *Our goal is to provide those tools, supported by theory that is easily understood and easy to implement.*

Some measurement processes may reasonably have Gaussian errors, but in the following common scenarios we have hard tolerance bounds:

- The digital two-digit readout is accurate (to within its round-off error)
- The error is less than 1% within the measurement range.
- The accuracy is ± 1 for values up to 20, then ± 5 up to 100, then ± 10 until top of the measurement range.

Such *interval inaccuracy* may come from round-off, other quantization or binning, hysteresis, or mechanical slack, and the quality control of instruments can impose error bounds. Given a measured value, we know an interval that contains the true value.

With actually stochastic measurements it may be difficult or impossible to infer their joint probability distribution. False assumptions made for convenience about the joint distribution, such as 'unbiased and symmetric', can dramatically mislead. If the inaccuracies are bounded, our approach can provide a safe and doable though pessimistic alternative without distribution assumptions.

This motivates our deterministic approach using interval inaccuracy.

Despite efforts [1] made in *interval analysis*, computations with numeric intervals remain unpopular. With only linear dependency on the unknowns, large numbers of inequalities or interval bounds can be combined to a linear program (LP) and solved numerically. For predictive models, the unknowns are the model parameters that need to be identified, and we focus on linear regression.

Many *compressed sensing* algorithms target the sparsest approximate solution to an underdetermined linear system. Some norm of the remainder (approximation error) must be small. There is currently little theory on inaccurate dictionary or sensing matrix, which corresponds to errors-in-variables in model identification. For an interval inaccurate dictionary our LP formulations provide sparse solutions, whose sparsity

may be further improved along the lines of Candes *et al.* [2].

Both L1-norm and sup-norm are appropriate error norms for LP. The nonlinear Euclidean norm emerges in statistics from Gaussian noise and maximum likelihood, which we avoid in our approximation theory approach that fits LP.

The theory section is self-contained, and describes the notation as well as basic results for linear interval systems. These essentials are needed to reach the main results, and the Discussion points to other related prior work. The elementary presentation does not require advanced concepts from robust convex optimization.

We will show that robustness induces sparsity/parsimony, provide means to pursue sparsity further, and formulate jointly sparse modeling as an LP. The L1 approximation provides solutions less sensitive to outliers, and while L1 regression has been widely practiced, our contribution past prior approaches is the treatment of errors-in-variables. The inclusion of redundant features in the model by regularization, in a fashion analogous to the Elastic Net, is also novel.

LP solvers are widely accessible, and enable robust small sample regression. Any common spreadsheet program includes an LP solver for at least up to 200 variables, while more dedicated LP implementations handle thousands of variables. Numerical toy examples are shown to illustrate the practical usage, and to provide insights essential to modelers.

2 Theory

2.1 Interval matrices

In linear regression, the model inputs or independent variables have their values collected to the design matrix A, while the targeted outputs are approximated by a linear model $b \approx Ax$ whose model parameters x need to be identified. For controlled experiments, the experimental design determines A, and a column of ones is added for the regression constant. In our case the inaccurate measurements in A are given as intervals. An interval matrix $[A] = [A_{min}, A_{max}]$ is defined by the matrices with its elementwise lower and upper bounds. A representative or realization $\tilde{A} \in [A]$ is an ordinary numeric matrix that conforms to these bounds $A_{min} \leq \tilde{A} \leq A_{max}$. Inequalities between vectors or matrices hold element by element. Alternatively $[A]$ can be described by its center and nonnegative radius as $[A] = [\check{A} \pm rad(A)]$, where $\check{A} = (A_{min} + A_{max})/2$ and $rad(A) = (A_{max} - A_{min})/2$.

Considering an interval matrix as a set of matrices, arithmetic operations are defined by including all possible results with that set. When an interval matrix multiplies an ordinary vector, the result is an interval vector $[A]x = [\check{A}x \pm rad(A)\,|x|]$, and its elements can independently of others get any values in their ranges. The absolute value is applied elementwise in $|x|$.

Similar notation is used with vectors and also scalars. For scalar intervals $[\alpha] \subseteq [\beta]$ if, and only if

$$\left|\check{\alpha} - \check{\beta}\right| \leq rad(\beta) - rad(\alpha)$$

This holds component by component when one interval vector is included (as a subset) in another, and is shortly needed.

2.2 Tolerance solutions as the approach to robustness

Let us now require that $[A]x \subseteq [b]$. Then the identified x gets mapped close to $\check{b}$ regardless of which realization from the set $[A]$ is selected. For an experimentalist, a noisy replication of the training data will give closely similar model predictions, so the model can indeed describe the cases. Combining the two basic results given,

$$\left|\check{A}x - \check{b}\right| + rad(A)\,|x| \leq rad(b)$$

When the output tolerance $rad(b)$ is given, this inequality defines so-called *tolerance solutions* [3; 4]. We will instead minimize $rad(b)$ in the spirit of Chebyshev approximation.

2.3 Sup-norm approximation program

A linear program that minimizes the **sup-norm** of $rad(b)$ is

$$\begin{array}{rcl} & \text{minimize } s & \\ x_+ \geq 0,\ x_- \geq 0,\ s \geq 0 & & \\ A_{max}x_+ - A_{min}x_- & \leq & b + se \\ A_{min}x_+ - A_{max}x_- & \geq & b - se \end{array}$$

where $e = (1, \cdots, 1)^T$ has all elements equal to one. The optimal solution is $x^* = x_+ - x_-$.

This is an ordinary convex LP despite the absolute values in the inequality system. In general, absolute value inequalities have a connection to complementarity problems [5], and easily give NP hard problems where, for example,

each orthant needs to be dealt with separately. This happens if one pursues the united solution set from Oettli-Prager theorem, given in the Discussion.

The tolerance parameter s is of practical *diagnostic* significance: it is the maximum deviation from the targeted output across the cases and the matrix realizations. A small tolerance relative to the important differences between elements of b indicates successful modeling, while a large tolerance indicates failure in model fitting. *It is better to know of this failure, and not pass a poor model on to fail in validation experiments.*

The sup-norm is sensitive to outliers or bad data points. A single case with a 'noise spike' can mainly determine the sup-norm, so the better data points are effectively masked during parameter identification.

2.4 L1-norm approximation program

The L1-norm is the sum of elementwise absolute values, also called the *city block distance* or the *Manhattan norm*. It gives robustness against outliers, as these can not mask the effects from other cases. The disadvantage is in interpretation of results: this norm depends on all of the cases, so it does not give a clear sharp bound on an individual case of interest, while the sup-norm does.

The **L1-norm** of minimum output tolerance $rad(b)$ is, from straightforward algebra,

$$\left\|\check{A}x - \check{b}\right\|_1 + \|rad(A)\,|x|\,\|_1$$

This is minimized by the LP formulation

$$\begin{aligned} \text{minimize } & e^T u + v^T(x_+ + x_-) \\ & x_+ \geq 0,\ x_- \geq 0,\ u \geq 0 \\ & \check{A}(x_+ - x_-) - b \ \leq \ u \\ & b - \check{A}(x_+ - x_-) \ \leq \ u \end{aligned}$$

where the constant vector $v^T = e^T rad(A)$ should be pre-computed, and e again has all elements equal to one. The minimum recovered is a tight L1-bound for the approximation error in $b \approx \tilde{A}x$ that any matrix realization $\tilde{A} \in [A]$ will conform to. The optimal identified parameters are again $x^* = x_+ - x_-$.

2.5 Control of sparsity and connection to Compressed Sensing

If we increase, say, the values in the first column of $rad(A)$, eventually the rate of worsening of the term $rad(A)\,|x|$ on moving the first component in x away from zero will dominate over the rate of improvement in the term $\left|\check{A}x - \check{b}\right|$. Then the optimal solution must have its first component stay at zero. So by increasing the matrix inaccuracy artificially, we can control the sparsity of these solutions.

We introduce the sparsity controlling parameter $t \geq 0$, and substitute $rad(A) \leftarrow t\ rad(A)$. For $t = 0$ this gives an accurate system, with no errors-in-variables, leading to ordinary sup-norm or L1-norm regression. For $t = 1$ we have the real inaccuracy from our metadata on the feature value determinations, while an increasing $t > 1$ will induce further sparsification until all inaccurate features are eliminated. This is analogous to the Lasso where the L1-norm of the solution is penalized, and increasing the penalty weight gives sparser solutions. After such feature selection, parameter re-identification with $t = 1$, with only these features, gives the most robust model for the real inaccuracy with the select features.

For an accurate feature the column in $rad(A)$ is null. A numerical toy example will later show that also such features may be dropped as irrelevant or unnecessary.

The metadata on inaccuracy enforces some level of sparsity or parsimony of the model. Robustness is a clear concept, here characterized by the matrix inaccuracy and the tolerance parameter s, and is of value to users of an identified model as a performance guarantee; parsimony is a secondary consequence, though it may affect the measurement costs on using the predictor.

Statistical methods for determining the proper level of model parsimony, such as Akaike or Bayes information criteria (AIC and BIC), are not transparent and do not match the case of errors-in-variables. In contrast our approximation theory approach is easy to understand, apply, and interpret. The two approaches fundamentally differ in their assumptions.

The connection to common *ad hoc* pursuit of sparsity by L1-regularization emerges with uniform matrix uncertainty $rad(A) = r\,ee^T$, where the two e vectors of ones need not have the same dimension (the matrix needs not be square). Then $rad(A)\,|x| = r\,e\,\|x\|_1$.

For sup-norm approximation, uniform matrix uncertainty gives

$$x^* = \mathrm{a}rgmin_x \left\|\check{A}x - \check{b}\right\|_\infty + r\,\|x\|_1$$

while for L1-norm approximation the regularization term scales differently

$$x^* = \mathrm{a}rgmin_x \left\| \check{A}x - \check{b} \right\|_1 + rn \left\| x \right\|_1$$

where n is the number of cases (rows in matrix A). The familiar penalty or regularization term popularized by Lasso emerged.

Such regularizer terms are commonly introduced as convex surrogates for support size, but here they emerged naturally from matrix inaccuracy.

The support of a vector is the index set of its nonzero entries. Support size is the count of those entries.

The connection to compressed sensing is now evident. Our more general regularizer, also inducing sparsity, uses a general nonnegative matrix $rad(A)$ to weigh the vector $|x|$, before application of a norm to create the goal function for optimization.

L1-regularization will undoubtedly remain popular, because often public data sets do not include $rad(A)$ or metadata that helps generate it. In this sense, Lasso –type L1-regularization is 'based on ignorance', and that lack of metadata is an indication it was considered unimportant because of inability to use it. This corroborates the relevance of current work.

2.6 Jointly sparse solutions

The LP formulations naturally deal with an excess of variables, and select the best basis from them. The basis has n features, equal to the number of rows in matrix A, limiting the number of nonzero coefficients. This enables compressed sensing –type applications where sparse solutions are sought for underdetermined systems. We have also shown how the sparsity can be driven further with the parameter $t \geq 0$.

Consider the jointly sparse solution of two systems of equations $b \approx Ax$ and $c \approx Ay$. We want the nonzero components of the two solutions to share the same locations by index, but their actual numeric values have no other mutual restraints.

One could employ binary 0/1 variables, indicating whether a particular feature is used or not. However, this leads to branch-and-bound techniques that are computationally prohibitively heavy even for small scale problems.

We instead use a large constant, say $M = 10^5$, and constrain $|y| \leq M\,|x|$. Then the support of y is included in the support of x, without much constraining the numeric values with M large.

Generally, the LP formulations for each of the linear approximation tasks are concatenated into one large formulation. In sup-norm approximation the same tolerance s is shared between the appropriately scaled related learning tasks. Now add to this formulation the solution vector bounds with a large M, so that one problem (the x) guides the rest (the y's) in their support selection.

With this large M approach for joint feature selection, the level of joint sparsity can be controlled with $t > 0$, while the approximation error can be measured by either sup-norm or L1-norm.

2.7 Imitating the Elastic Net

When small-scale pilot experiments are run with the intent to follow up with a more expensive larger set of experiments, the purpose of the pilot is to reduce costs by limiting the set of measurements to selected features, as well as to contribute to the model-based design of the follow-up. For such applications, it is important to retain some safety margin and not make the most fragile bet. Taking the sparsest or most parsimonious predictor saves the most measurement cost, but is also the most fragile bet — losing any of the features gets below the most parsimonious that was still useful.

Loss of observability is natural on scale-up. Bench-scale and pilot equipment are designed for observability, while industry scale is designed for production. Similarly pre-clinical experiments are by their nature more observable than live clinical patients, who object to or forbid invasive observations.

A margin of safety against this loss means retaining redundant features from the pilot data. This is exactly how the Elastic Net transforms the Lasso. Effectively, it progressively penalizes large model parameters and reduces heavy betting on any single predictive feature. Useful predictive features get their 'fair shares' in predictive roles.

The Elastic Net, despite its many successes, is not designed to deal with errors-in-variables. It does include the L1-penalty, which we know could emerge from uniform matrix inaccuracy, but this and the square of Euclidean norm of the parameter vector are used as *ad hoc* regularizations with indeterminate weights. We have

seen that the matrix inaccuracy induced regularization weight changes depending on the selected approximation error norm, and want to here develop an LP-based alternative that properly matches the inaccuracy.

The least squares solution of $b \approx Ax$ is analytically given by $x = (A^T A)^{-1} A^T b$ when the inverse exists, and when it does not one pads the equation system to $\begin{pmatrix} b \\ 0 \end{pmatrix} \approx \begin{pmatrix} A \\ \lambda I \end{pmatrix} x$ for which the same formula gives the analytical solution $x = (A^T A + \lambda^2 I)^{-1} A^T b$ that is always well defined for $\lambda > 0$. The least squares problem that has just been solved can be expanded to

$$\|b - Ax\|_2^2 + \lambda^2 \|x\|_2^2 = \text{minimum}$$

where the Tikhonov regularization is explicit. We view the Elastic Net as adding L1-regularization to this, to also pursue sparsity. With our robust solution approach sparsity is automatically induced, so we only imitate the steps before adding the L1-term.

Using the padded equation system with an interval matrix $[A]$, and the sup-norm for approximation error

$$\begin{aligned} \left|\check{A}x - \check{b}\right| + rad(A)\,|x| &\leq se \\ \|x\|_\infty &\leq s/\lambda \end{aligned}$$

which for uniform inaccuracy $rad(A) = ree^T$ simplifies to

$$\begin{aligned} \left\|\check{A}x - \check{b}\right\|_\infty + r\,\|x\|_1 &\leq s \\ \|x\|_\infty &\leq s/\lambda \end{aligned}$$

For a practical LP formulation, the second inequality expands to p inequalities, as many as there are components in x. This can hamper the numeric computations, especially if the simplex method is used.

Iteratively, start with solving only the first inequality. Impose the second only on those components that have violated it in any solution so far, and re-solve. We are done when no further violation occurs.

Here, λ is an *ad hoc* regularization parameter, and we do not know of any accepted quantification of solution redundancy. The robustness of the solution can not possibly improve from adding constraints with $\lambda > 0$, and the worsening tolerance s may give a hint on what values of λ are reasonable. On the other hand, the purpose of the approach is to keep an excess of features, as those deemed utterly useless will not be measured at all in the next experiments designed. The approximation quality of the constrained redundant solution is not necessarily relevant. Thus we are incapable of providing a complete solution, and only provide some useful elements to approach it.

An alternative approach could be to find a primary non-redundant model, then in a leave-one-out pattern with the selected features try to find redundant optimal replacements for each primary feature in turn. If the tolerance dramatically increases from leaving a feature out, then this feature is particularly important to keep. However, there is no guarantee that that the remaining features from the primary model will be re-used; instead the new models may select entirely non-overlapping feature sets, when there is an over-abundance of candidate features available (like often in genomics).

Despite its potential flaws, the above regularization approach enables adding only some small margin of redundancy, based on feature count, as opposed to a large multiple in size of the most robust feature set.

3 A Numerical Toy Example

We illustrate the use of robust regression by LP, to show how mixed types of inaccuracies are handled, and how the user choices influence the solutions.

Let the system matrix and its radius be

$$\check{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix},\ rad(A) = \begin{pmatrix} 0.2 & 0 & 0.5 \\ 0.2 & 0 & 0.5 \\ 0.2 & 0 & 1.5 \end{pmatrix}$$

showing uniform inaccuracy for the first feature, accurate values for the second, and proportional inaccuracy for the third.

For $b = (0, 1, 2)^T$ the accurate solution is $x = (\frac{-2}{3}, \frac{1}{3}, \frac{2}{3})^T$, while the most robust sup-norm approximation to b is given by $x^*_{sup} = (0, 0, \frac{2}{3})^T$. It seems counterintuitive that only the least accurate third feature is selected. However, with L1 approximation the solution is $x^*_{L1} = (0, 1, 0)^T$ — the inaccurate features are dropped, the third case is poorly reproduced as an outlier, but the first and second cases are accurately reproduced. Both robust solutions are sparse, as expected.

Sup-norm regularization of the solution forces the dropped features back in, and limiting the

absolute component sizes to 0.5 gives $\tilde{x}_{L1} = (\frac{-1}{2}, \frac{1}{2}, \frac{1}{2})$. The approximation error for this solution seems better than without regularization, but this is only at the center of the interval matrix; the worst case error has increased.

In this case the system matrix is invertible, and a Total Least Squares (TLS) solution ignores the inaccuracy completely because a fully accurate solution is available. In other words, the known inaccuracies of inputs have no effect whatsoever on a TLS solution. This characteristic of TLS may give a high weight to an inaccurate feature, so the model output has high tolerance or uncertainty. That uncertainty is minimized by robust model fitting, but completely ignored by TLS.

By scaling up the inaccuracy in this example case, the sup-norm approximation is driven to a null vector (all features get eliminated), while the L1 approximation keeps the accurate feature with weight 1. TLS would remain unaffected.

This example has highlighted a dramatic difference to Total Least Squares, and the outlier detection with L1 approximation. When the inaccuracies in $rad(A)$ vary widely between rows, a case with high inaccuracies can determine the sup-norm alone.

Targeting another output $c = (0, 1, 0)^T \approx Ay$ with a jointly sparse solution, with sup-norm approximation error, gives unchanged x^*_{sup} and the new solution $y^*_{sup,joint} = (0, 0, 0)^T$. The accurate separate solution would be $y = (0, 1, 0)^T$, but the middle component is forced to zero as it is not in the support of $x^*_{sup} = (0, 0, \frac{2}{3})^T$ and does not improve the sup-norm of the residual.

This tiny MMV example shows how the problems are coupled by their scaling and the error norm that is minimized — the large M only forces similar support, and is not the only coupling on solving an MMV problem. On setting up an MMV problem the targeted outputs must be scaled so that their approximation errors are of similar order of magnitude.

4 Discussion

The theory demonstrates how interval inaccuracy in the design (or dictionary) matrix is dealt with by requiring robust tolerance solutions. Identifying the model parameters that minimize this tolerance turned out to be a single convex and sparsifying linear program. This is surprising since the governing inequalities are absolute value equations.

An earlier approach by Rosenbaum and Tsybakov [6] uses in contrast the Oettli-Prager solution set (see also Beaumont and Philippe [3], and Polyak and Nazin [7]), which is *non-convex unless the solutions are restricted to a single quadrant*. They impose *ad hoc* regularization with L1-norm to find the sparsest vector in this solution set, and in their numerical examples restrict the solutions to the nonnegative quadrant. The equations provided are *limited to the case of uniform matrix uncertainty*, while the general Oettli-Prager theorem would allow general matrix uncertainties with any nonnegative $rad(A)$.

Our concise derivation of inequalities for tolerance solutions is adapted from Mayer [8], while the book by Fiedler *et al.* [4] gives more complicated and comprehensive theory of linear interval systems.

Following Mayer's approach, also the Oettli-Prager theorem is quick to derive. For scalar intervals $[\alpha] \cap [\beta] \neq \emptyset$ if, and only if

$$\left|\check{\alpha} - \check{\beta}\right| \leq rad(\alpha) + rad(\beta)$$

which leads to the united solution set (or possibilistic solutions) of an inaccurate linear equation system

$$\left|\check{A}x - \check{b}\right| \leq rad(A)\,|x| + rad(b)$$

This result is the Oettli-Prager theorem [4].

Given an accurate b, solution x is possible for some matrix realization, $b \in [A]\,x$, exactly when

$$\left|\check{A}x - \check{b}\right| - rad(A)\,|x| \leq 0$$

which under uniform matrix uncertainty determines the possibilistic solution set by

$$\left\|\check{A}x - \check{b}\right\|_\infty - r\,\|x\|_1 \leq 0$$

where we have the difference of two convex terms in x. Rosenbaum and Tsybakov used this to define the feasible parameter vectors x, and picked a sparse possibilistic solution by L1-norm minimization. When restricted to a single quadrant, the convex L1-norm is just a linear function, so solution as an LP becomes possible. Our tolerance solutions have a single sign change from Oettli-Prager, and provide a convex optimization problem without need for quadrant restriction.

With increased matrix inaccuracy the set of possible matrix realizations grows, and so does the set of possibilistic solutions. As the feasible set grows, a better optimum becomes possible. In this way, increased inaccuracy makes the problem easier, in the approach of Rosenbaum and Tsybakov. Our approach behaves the opposite way, with increased inaccuracy giving a heavier penalty and 'making parameter identification more difficult'. This seems reasonable and preferable.

We showed that minimizing the tolerance, whether in terms of sup-norm or L1-norm, leads to sparse solutions, and increased matrix inaccuracy induces higher sparsity (smaller support size). The level of model parsimony is simply a side-effect from robustness. In compressed sensing sparsity can be a value in and of itself, and faking higher inaccuracy allows us to force a higher level of sparsity. This will come at the cost of a higher output tolerance, in other words the model becomes less robust against inaccurately replicated training data. A sparse model may save instrumentation or other observation costs, and be therefore desirable.

Our results are in alignment with the robust regression results on Xu *et al.* [9], who focus on Euclidean approximation error. Also in their cases, robustness against noisy model inputs induces sparsity of the solutions. The benefit of our approach is doability with common LP solvers.

The L1-norm as approximation error will be less sensitive to outliers, compared to the sup-norm (or the Euclidean norm). While the sup-norm in the spirit of Chebyshev approximation will likely be the preferred choice of most users, we provided the LP formulation for also L1-norm. This has a connection to the problem where the data are otherwise accurate but some cases are disturbed by error spikes (or occasional transmission errors), discussed by Candes, Wakin, and Boyd [2]. In other words, the targeted output is sparsely perturbed, and this perturbation is found by L1-norm bounding it; befitting the capability of L1-minimization to recover sparse unknowns. Their reweighted L1 minimization provides an LP approach to extreme solution sparsity, that could be adapted to the more general sparsity-inducing regularization that emerged in our tolerance approach.

The joint sparsity or MMV-problem (for Multiple Measurement Variables) has apparently not been solved before with an LP approach. The works of Cotter *et al.* [10], Lee and Bresler [11], and Mishali and Eldar [12], are representative highlights. Our LP approach provides a special alternative, likely far behind the fastest and memory-limited to smallish problems, but with optimal robustness.

For modeling practitioners involved in sequentially designed experiments, the capability to do something analogous to the Elastic Net is important. Limiting measured variables to the smallest subset that seems to perform in preliminary experiments is risky, and a methodical approach to more lenient feature selection is valuable. For small problems all the sup-norm constraints can be included in one LP solution, while for larger problems iterative stepwise enforcement of the sup-norm limit will converge by exhaustion.

Sup-norm regularization is rare in the context of compressed sensing. Donoho, Maleki, and Montanari [13] show its use without explanation of motivation or background. Hindi and Boyd [14] have devised various convex optimization approaches to robust solutions with inaccurate data, but the norm bounds used seem distant from practice. Their Theorem 7 shows regularization with the sup-norm of the solution vector. In our case the sup-norm seems a trivial addition, once we consider giving an uncertain solution out (to a customer) as making a risky bet, and we don't want to bet too much on any single feature. An analogy to the Elastic Net vs. Lasso was established, by using the derivation of Tikhonov regularized solution formulae for least squares problems.

5 Conclusions

With common interval inaccuracy in the predictive features, robust solutions that minimize approximation tolerance generally lead to sparse solutions, and model parsimony is a consequence of optimal robustness. The widely practiced *ad hoc* L1-regularization, to promote solution sparsity, is just a special case corresponding to uniform matrix uncertainty.

Linear programs were provided for both sup-norm and L1-norm bounds on the approximation error. The former is for reliable though inaccurate data, while the latter is better with spike noise in the modeled data making some cases

outliers. Also the jointly sparse solution of a multitask learning problem was formulated as an LP.

Because solutions to linear programs can be found with certificates of optimality, these formulations can reliably help benchmark new numerical approaches under development. By this means they can support research on algorithms for compressed sensing.

Especially for stepwise experimental design, where pilot experiments decide which features are measured in later experiments, it is important to leave some redundancy in the feature selection, as the sparsest feature set has no fault tolerance remaining. A size bound for the identified parameters, added to the linear programs above, addresses this concern.

The array of formulations provided serves as a collection of easily implemented recipes for those practicing modeling with significantly inaccurate input features (errors-in-variables). Common spreadsheet programs can be used to solve linear programs with at least up to 200 unknowns, and both free and commercial software packages are available that are designed to deal with thousands of unknowns. The LP solvers can be considered a well-developed reliable computational technology. The recipes provided are therefore practically usable for anyone with access to a computer.

Acknowledgment

The author gratefully acknowledges general support by the Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani, Thailand.

References

[1] R. E. Moore. *Interval analysis*. Prentice-Hall, 1966.

[2] E. J Candes, M. B Wakin, and S. P Boyd. Enhancing sparsity by reweighted l1 minimization. *J. Fourier Anal. Appl.*, 14:877–905, 2007.

[3] O. Beaumont and B. Philippe. Linear interval tolerance problem and linear programming techniques. *Reliable computing*, 7(6):433–447, 2001.

[4] M. Fiedler, J. Nedoma, J. Ramik, J. Rohn, and K. Zimmermann. *Linear Optimization Problems with Inexact Data*. Springer Science+Business Media, Inc., 2006.

[5] O.L. Mangasarian. Absolute value programming. *Computational Optimization and Applications*, 36(1):43–53, 2007.

[6] M. Rosenbaum and A. B. Tsybakov. Sparse recovery under matrix uncertainty. *Annals of Statistics*, 38(5):2620–2651, December 2010.

[7] B.T. Polyak and S.A. Nazin. Interval solutions for interval algebraic equations. *Mathematics and Computers in Simulation*, 66:207–217, 2004.

[8] G. Mayer. On regular and singular interval systems. *Journal of computational and applied mathematics*, 199(2):220–228, 2007.

[9] H. Xu, C. Caramanis, and S. Mannor. Robust regression and lasso. In *Advances in Neural Information Processing Systems 21*, pages 1801—1808, 2009.

[10] S.F. Cotter, B.D. Rao, K. Engan, and K. Kreutz-Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 53(7):2477–2488, July 2005.

[11] K. Lee and Y. Bresler. Subspace-Augmented MUSIC for joint sparse recovery with any rank. In *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2010 IEEE*, pages 205–208, April 2010.

[12] M. Mishali and Y.C. Eldar. Reduce and boost: Recovering arbitrary sets of jointly sparse vectors. *IEEE Transactions on Signal Processing*, 56(10):4692–4702, October 2008.

[13] D. L Donoho, A. Maleki, and A. Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914, 2009.

[14] H. Hindi and S. Boyd. Robust solutions to l1, l2, and l-infinity uncertain linear approximation problems using convex optimization. In *Proceedings of the American Control Conference*, volume 6, pages 3487–3491, June 1998.

The extracting of shirt composition using Canny edge detection with close contour

Lanjakorn Sewata Salil Boonbrahm

School of Informatics
Walailak University
Nakorn Si Thammarat, Thailand. 80160
lsewata@me.com, salil.boonbrahm@gmail.com

Abstract

In this paper, we present the procedure that extracting image of a shirt into several components in order to reuse this component in the future. Canny edge detection is used to identify edge of the image and close contour is then used to identify components of the shirt where the shirt library is used as reference together with intercorrelation and autocorrelation functions.
Keywords: Image Segmentation, Close Contour, Edge Detection, Shirt Extraction.

1 INTRODUCTION

The objective of this research is to provide tool to manipulate with shirt images that will be inputted into the system. With this tool, we will be able to enhance the image of shirt to reduce noise, to recognize the shirt edge and then extract the shirt image into different parts such as neck, right sleeve, left sleeve, and so on.

The first part of the article contains reviews of definitions and algorithms in image enhancement, edge detection and image segmentation. These algorithms are applied to use on our problem domain where we try to segment the shirt image. The second part is the proposed model that designed to extract the shirt image by using shirt library and close contour of small elements of the shirt.

2 RELATED WORKS

2.1 Image Enhancement

In image enhancement, the goal is to accentuate certain image features for subsequent analysis or for image display. Examples include contrast and edge enhancement, pseudo coloring, noise filtering, sharpening and magnifying. Image enhancement is useful in feature extraction, image analysis and visual information display. The enhancement process itself does not increase the inherent information display in the data. It simply emphasizes certain specified image characteristics. Enhancement algorithms are generally interactive and application dependent.

Image enhancement techniques such as histogram equalization, contrast stretching, grey level into another grey level by a predetermined transformation. An example is the histogram-equalization method, where the input levels are matched so that the output grey level distribution is uniform. This has been found to be powerful method of enhancement of low contrast images. Other enhancement techniques perform local neighborhood operations as in convolution; transform operations as the discrete Fourier transforms.

The best known order-statistics filter is the median filter, which replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel

$$\hat{I}(x,y) = \underset{j=1,2\ldots,9}{Median}\{I(a_j)\} \qquad (1)$$

The original value of the pixel is included in the computation of the median. Median filters are quite popular because, for certain types of random noise they provide excellent noise reduction capabilities, with considerably less blurring than linear smoothing filters of similar size [5][6].

2.2 Image Segmentation

Image segmentation, one of the fundamental problems of early computer vision, has been intensively studied in the past. The existing automatic image segmentation techniques can be classified into three approaches, namely, 1) thresholding techniques, 2) boundary-based methods, and 3) region-based methods.

1) Thresholding techniques are based on the assumption that adjacent pixels whose value (grey level, color value, texture etc.) lies within a certain range belong to the same class. Thresholding techniques can obtain good segmentation of images that include only two opposite components. Since these techniques neglect all the spatial relationship information of the images, they are inefficient for images that blur at object boundaries, or for multiple image component segmentation.[1]

2) Boundary-based techniques use the assumption that pixel values change rapidly at the boundary between two regions. Edge detectors used in these techniques can be simple ones such as the Sobel or Roberts operators, ormore complex ones such as the Canny operator. The output of most existing edge detectors can only provide candidates for the region boundaries, because these obtained color edges are normally discontinuous or over-detected. However, the actual region boundaries should be closed curves. Therefore, some post-procedures, such as edge tracking, gap filling, smoothing, and thinning, should be performed to obtain the closed region boundaries. All of these post-procedures are very time-consuming; converting the edge candidates to the region boundaries is thus not an easy task. The time-consuming post-procedures can be avoided by integrating the results of the boundary-based approach and those of the region-based approach.[1]

3) Region-based techniques rely on the assumption that adjacent pixels in the same region have similar visual features such as grey level, color value, or texture. A well-known technique of this approach is split and merge. Obviously, the performance of this approach largely depends on the selected homogeneity criterion. Instead of tuning homogeneity parameters, the seeded region growing (SRG) technique is controlled by a number of initial seeds . Given the seeds, SRG tries to find an accurate segmentation of images into regions with the property that each connected component of a region meets exactly one of the seeds. Moreover, high-level knowledge of the image components can be exploited through the choice of seeds. This property is very attractive for semantic object extraction toward content-based image database applications. However, SRG suffers from another problem: how to select the initial seeds automatically for providing more accurate segmentation of images.

2.3 Edge Detection

Edge detection is one of the first steps in many pattern recognition and computer vision pipelines. It is an image processing technique that aims at the identification of sharp changes or discontinuities in the image brightness within neighbor regions on the image plane.

The determine threshold value is an important factor in finding the edge because threshold values that is less or too much affect the performance edge detection image and It also helps in reducing the noise of the image in Fig 1.

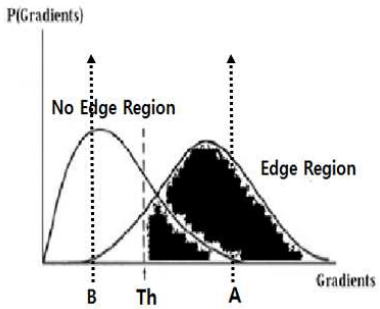


Figure 1 the determine threshold value [2]

In Fig 1, when the threshold values move to position A, the border area will then reduce that it can reduce the noise of the image. It move to position B is edge region will be increase that the noise of the image is increased accordingly. This research is choose to use the threshold in position B. [1],[2]

The Canny Edge Detection Algorithm is an optimal edge detector having good localization property, such that the marked edges are as close as possible to the edges in the real image and minimal response, so that one image edge should only be marked once, and only when there is strong evidence that it is really an important feature.

The first step on the algorithm is the Canny Edge Detection Algorithm, which performs noise removal, gradient detection and edge strength analysis respectively.

First of all, it is necessary to apply a Gaussian Smoothing to I, expressed as the convolution J = I*g of I with a Gaussian kernel g.

After this, the following steps must be done for each pixel (i, j) in J:

a) Compute the gradient components J_x and J_y

b) Estimate the edge strength E_s in EQ 2.

$$Es(i, j) = \sqrt{Jx(i, j)^2 + Jy(i, j)^2} \quad (2)$$

c) Estimate the orientation of the edge normal E_0 in EQ 3.

$$Eo(i, j) = a\tan\frac{Jy}{Jx} \quad (3)$$

The output is a strength image Es, formed by the values Es(i, j) and an orientation image Eo, formed by the values of Eo(i, j).

After that, the non-maxima suppression Algorithm is applied to the strength image, considering the four directions identified by 0°, 45°, 90° and 135°, in order to remove non-maxima points in the Es image and produce a new image In.

For each pixel (i,j) in Es, it is necessary to find the direction Dk, which better approximates the direction Eo(i, j), which is the edge normal. If Es(i, j) is smaller than at least one of its two neighbors along Dk, then In(i, j) = 0 otherwise In(i,j) = es(i,j). The output image In(i,j) consists of points belonging to thinned edges (i.e. Es after suppressing non maxima points).

The image produced by the previous step still contains local maxima originated by noise. In order to get rid of such local maxima, the next and last step of the process consists in using the Hysteresis Threshold Algorithm, which uses a minimum and maximum thresholding values t_l and t_h.

First, the next unvisited edge pixel, In(i, j), such that In(i, j) > t_h is located. In the sequel, the algorithm starts from In(i, j), following the connected chains of local maxima in both directions perpendicular to the edge normal, as long as In(i,j) > t_l

In the next step, all visited points are marked and a list of the locations of all points in the connected contour found is saved. The output is a set of lists, each one describing the position of a connected contour in the image, as well as the strength and the orientation images, describing the properties of the edge points. [3],[4]

Figure 2 (1) Source Image
(2) Image results by canny algorithm [2]

3 Methodology

The methodology using in this research contain four processes which are image enhancement, edge detection, image segmentation and image extraction. In order to use these processes the input which is images or patterns of shirts are captured. Then image enhancement is used to reduce the noise of the image using median filter algorithm. Edge detection is indicates the position of the shirt using Canny edge detection algorithm. Images segmentation on the shirt edge used to calculate the area and shape of the edges, which is characterized by the appearance of a closed contour, which is to be compared with the shirt images in library to perform extraction elements picture as an shirt using Intercorrelation function (ICF) and Autocorrelation function (ACF). If ICF value and ACF value than specified value in the system assumes as elements of image in Fig 3

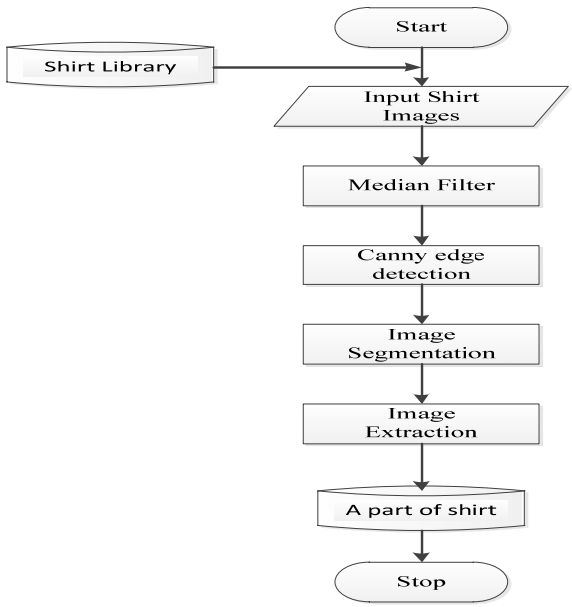


Figure 3 Shirt extraction proceduce.

3.1 Shirt Library

To use the shirt extraction procedure, the shirt library has to prepare first. Shirt library contains the images used for comparison with the close contour. These images are used in the process of image extracting

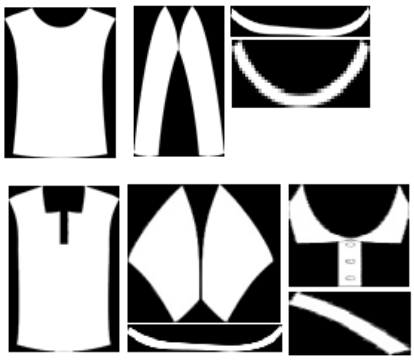


Figure 4 Example of images in library.

3.2 Image Segmentation

At first, we define such an object contour. The contour is a boundary of object, a population of points, separating object from a background. In systems of computer vision, some formats of coding of a contour are used - the codes of Freeman, two-dimensional coding, polygonal coding are most known. But all these formats of coding are not used in a contour analysis.

Instead, in a contour analysis the contour is encoded by the sequence consisting of complex numbers. On a contour, the point, which is called as starting point, is fixed. Then, the contour is scanned (is admissible - clockwise), and each vector of offset is noted by a complex number a+ib. where a - point offset on x axis, and b - offset on y axis. Offset is noted concerning the previous point in Fig 5.

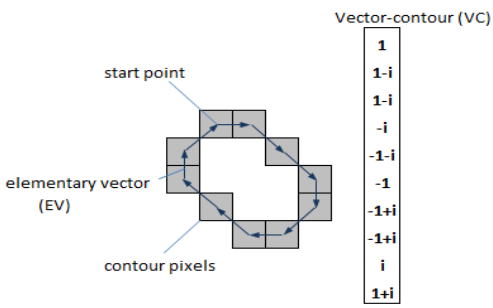


Figure 5 Contour Image

Owing to the physical nature of three-dimensional objects, their contours are always closed and cannot have self-intersection. It allows to define unambiguously a way of by-pass of a contour (to within a direction - on or counter-clockwise). The last vector of a contour always leads to the starting point.

Each vector of a contour we will name elementary vector (EV). And sequence of complex-valued numbers -vector-contour (VC).

Thus, vector-contour **σ** of length k can be designated as

$$\sigma = (\gamma_1, \gamma_2, \gamma_3, \cdots \cdot \gamma_{k-1}) \qquad (4)$$

Complex coding is close to two-dimensional coding where the contour is defined as a population of the EVs presented in the two-dimensional coordinates. But differences between operation of scalar product for vectors and for complex numbers are various.

The sum of an EV of a closed contour is equal to zero. It is trivial - as the elementary vectors result in starting point, their sum is equal to a zero-vector

$$\sigma = \sum_{n=0}^{k-1} (\gamma_n) \qquad (5)$$

where k - dimensionality of a VC, γ_n - n the elementary vector of contour **σ** is write pseudo code in Fig 6.

```
Function CloseContour
    Read Edge Images
    While (White Pixel)
        If (WhitePixel(x,y) = = Shift Right )
            a = 1
        ElseIf (WhitePixel(x,y) = = Shift Left )
            a = -1
        ElseIf (WhitePixel(x,y) = = Shift Up)
            b = 1
        ElseIf (WhitePixel(x,y) = = Shift Down )
            b = -1
        Else
            a= 0, b=0
        End
    VectorContour[i] = a+bi
    Next White Pixel
    If Summation (VectorContour[i]) == 0
        Output CloseContour Image
    End
End
```

Figure 6 Close contour Function

3.3 Image Extraction

Extraction composition algorithm is a comparison between a close contour with the image in shirt library. Using intercorrelation function

(ICF) and an autocorrelation function (ACF) can be designated as

$$ICF = max\left(\frac{\sigma(\mathbb{N})}{|\sigma||\mathbb{N}|}\right) \quad (6)$$

ICF is similar of shape between 0-1.
σ is Vector Contour of Close contour
$\mathbb{N}$ is Vector Contour of shirt template

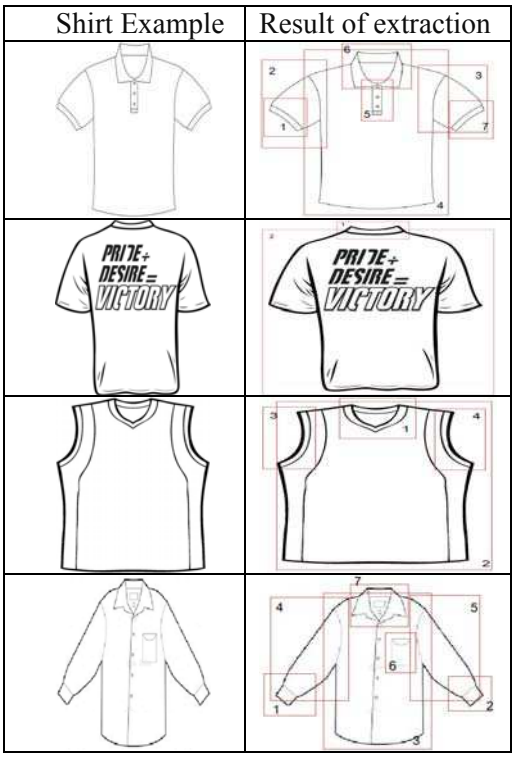
$$ACF = (\sigma, \mathbb{N}^m) \quad (7)$$

ACF is similar of length between 0-1.
σ is Vector Contour of Close contour
$\mathbb{N}$ is Vector Contour of shirt template

Which close contour is a part of shirt to be considered at the ICF and ACF calculated values should be greater than or equal to ICF and ACF assigned by the system.

4 Experimental Results

This research developed with Emgu CV library and use shirt of different 50 styles stored in formats JPEG and the size of the image is 1024 x 768 pixels defined by the ICF was 0.5 and the ACF was 0.85, with results as shown in Fig 7.



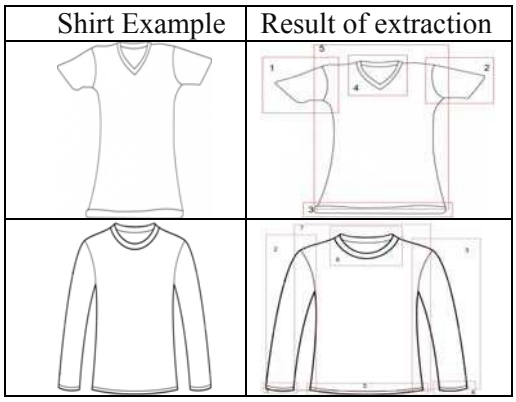


Figure 7 Result of shirts extract.

5 Conclusions

To obtain the complete and correct shirt extraction process, there are some factors that should take into consideration. Such factors are as follows:

1) The shirt library. The use of shirt library to compare the components of shirt image plays important role in identifying position of the component correctly and precisely.
2) The values of ICF and ACF. If setting the values of ICF and ACF too low, the precision in identifying the shirt components will be low.

The limitation of this research is that the input shirt images should have sharp and connected lines in order to identify each part or component of the shirt correctly. Further research will be focus on preprocess to enhance the shirt image that contained disconnected line.

References

[1] FAN, Jianping, et al. Automatic image segmentation by integrating color-edge extraction and seeded region growing. *Image Processing, IEEE Transactions on*, 2001, 10.10: 1454-1466.
[2] B.Poornima, Y.Ramadevi and T.Sridevi. Threshold Based Edge Detection Algorithm. *International Journal of Engineering and Technology*, Vol.3, 4-8, 2011.
[3] LEE, Tzu-Heng Henry. Edge detection analysis. Graduate institute of communication engineering, Taipei, Taiwan, ROC, 2007.
[4] ROQUE, Marcelo G., et al. Adapting the Sobel Edge Detector and Canny Edge Extractor for iPhone 3GS architecture. In: *17th Interna-*

tional Conference on Systems, Signals and Image ProcessingIWSSIP. 2010.
[5] GUPTA, Gajanand. Algorithm for Image Processing Using Improved Median Filter and Comparison of Mean, Median and Improved Median Filter. *IJSCE ISSN*, 2011, 2231-2307.
[6] Pei-Eng Ng and Kai-Kuang Ma, "A Switching Median Filter with BDND for Extremely Corrupted Images", IEEE Trans Image Processing, Vol. 15, No. 6, PP. 1506-1516, June 2006.

Digitally Generated Khmer Ornament for Khmer Architecture

Puthnith Var† Sokha Heng† Por Heng† Saovorak Khoy†

Pheakdey Nguonphan† Des Phal†

†Royal University of Phnom Penh, Phnom Penh, Cambodia
{var.puthnith, heng.sokha, heng.por, khoy.saovorak,
nguonphan.pheakdey, phaldes}@rupp.edu.kh

Abstract

Khmer ornament is really important for not only architecture but also Khmer civilization. Because the development of technology dramatically increases, computer becomes the main need for scientists, developers, engineers, designers, and architects. The ornament plays an important role in Khmer architectural design. Since there is no graphical art of the ornament, designers and architects face a difficulty in constructing the ornament again and again. To challenge the problem, we aim to aid them with programmatically generating the art whenever they need. Instead of storing each image of Khmer ornament, we generate the ornament based on a few scalable values or properties such as width and height of the art. In this paper, we experience on a Khmer ornament flower called *Phkar Chan* for our development.

Keywords: Khmer, Ornament, Architecture, Vector, Digitalization

1. Introduction

Khmer ornament or Kbach (in Khmer) is an important decorative element used extensively in objects and Khmer architecture throughout Cambodia [1]. Artists and architects create Kbach by using ornamental language, a method to create ornaments from basic forms to complex forms. To create complex ornaments, it takes long time despite the help from vector drawing or Computer Aid Design (CAD) software. So far, there is no method nor computer algorithms to generate module based Khmer ornament. Therefore, Artists or architects still need to draw them from scratch by using available tools in the software.

From a related work in [2], the author introduced module-based computer reconstruction methods, which make use of the structuring rule of the lotus motifs. With a similarity of characteristics of lotus motifs and Khmer ornament, this paper introduces similar methods to automatically generate module based Khmer ornaments from a basic form to a complex form.

2. Khmer Ornament Module

At the beginning of our work, we choose *Phkar Chan*, a flower art of Khmer ornament, to be our experience. The reason we choose it is because it is simple and it is commonly used in Khmer architecture and design. In this section, we discuss on the idea and how we put it into the implementation.

2.1. Idea

We want to generate Phkar Chan module without storing any point or data of it. Since the ornament module has to be drawn in square shape, we think the side of the square is the best parameter for the module. Therefore, we programmatically generate the Phkar Chan module based on the particular value.

2.2. Digitalization

There are two main components of Phkar Chan module: the petal and the heart. The petal component contains two small components can be seen in Fig. 1. The heart has its ring and petal. Furthermore, the module also has some other components such as the sepal.

Each component's location and size is set depended on the square's side parameter. In this basic of the module, we see that two half-petals is formed into a petal and the petal of the module is mirrored in opposite position to each

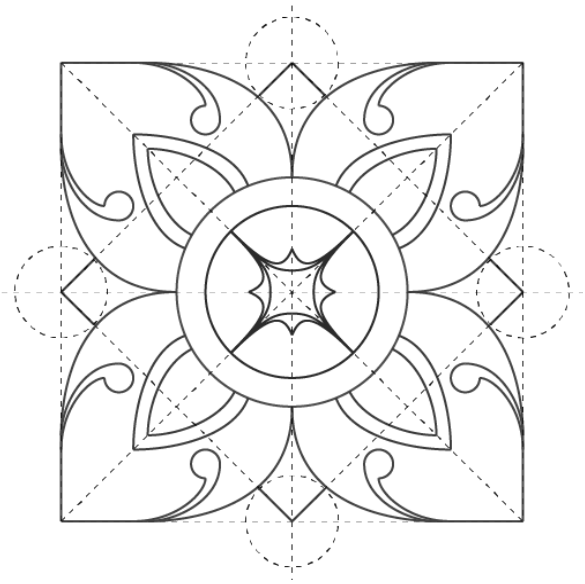


Figure 1. A sample of the *Phkar Chan* module

other. Therefore, we build on the half-petal component for the full set of the petal. Likely, the ovary of the module can have three and four petals depending on the style of the Phkar Chan module. Thus, to generate the module, we need the half-petal, the ovary's petal and ring, and the sepal components. In the complex module of Phkar Chan, each component is used repeatedly according to the pattern of Khmer architecture. Also some other components will be built if they are needed.

2.3. Terminology

Basically, each component is built using point and line. Those lines are drawn using many points. In simple way, straight line needs two points – start and end points. However, to draw a curve line, like Cubic Bezier curve, we need two additional points, which are the control points. The problem is how to find those points? Those points are found using the methods below. However, those points are related to each other based on the square's side parameter.

2.4. Methods

We use two different methods to find points. First method is called "*Grid-based Point Finding*". In this method, we construct each point based on gridding the square's side. In our experience, Big or simple is needed the grid up to the side divided by 4. Additionally, the grid's side is equals to the square's side divided by 128 if the component is complex or small compared to the square's side.

Another method is called "*Geometry-based Point Finding*". The points are found by intersecting lines and geometry's points. For example, we can find three points from equilateral triangle.

2.5. Experiment

In this experiment, we illustrate the ovary with three petals component using geometry-based point finding method.

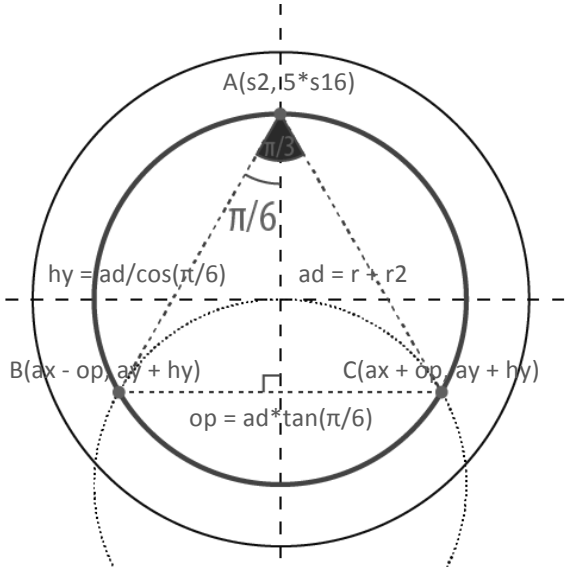


Figure 2. Three base points construction using *geometry-based point finding method* for the *ovary with three petals* component

Fig. 2 shows about constructing three base points using the methods. To draw the ovary, we use equilateral triangle to divide an inner circle into three portions. Point A can be found using grid-based point finding method. However, point B and C are found by geometry-based point finding method. To draw a Bezier curve between the three points, we need other six control points. To get a prefect ovary, we need sixteen points using our second method as shown in Fig. 3a. Fig. 3b and Fig. 3c show the uses of cubic and quadratic Bezier curves to draw the petals of the ovary. Finally, we get the beautiful ovary as shown in Fig. 3d. This proves that each Khmer ornament uses mathematic and nature to be drawn rather than it is just to be beautiful and meaningless.

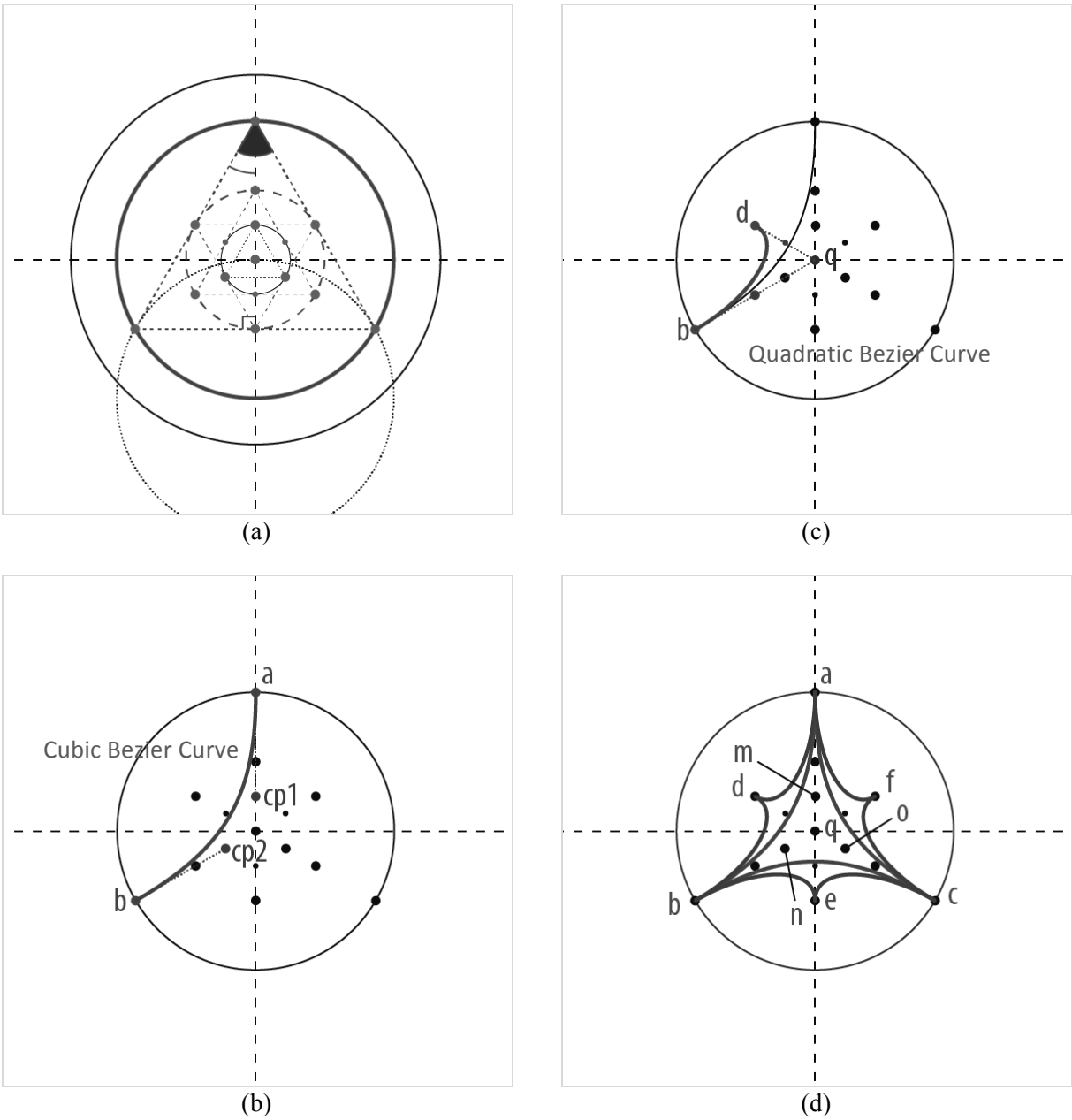


Figure 3. The process of illustrating the *ovary with three petals* component of *Phkar Chan* module

2.6. Modules

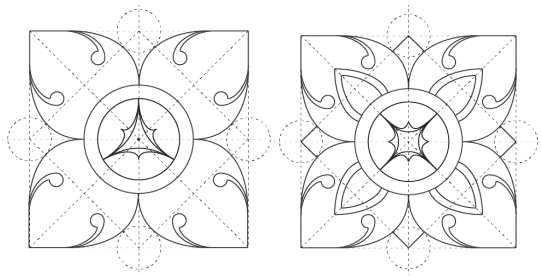


Figure 4. *Phkar Chan* modules (the left side is module A, the right side is module B)

We currently have two modules. Fig. 4 shows the module A that has only one petal and an ovary with 3 petals *and* the module B that has two petals, a sepal, and an ovary with 4 petals.

3. Conclusion

Pkhar Chan is programmatically generated based on a side of its frame. We use two different methods to find points to draw each component of the module. We research on the characteristic of the flower and draw it to the original form of

our ancient time. We apply mathematic to the design and make the work of drawing Khmer ornament become faster and better. We believe that this small piece of research will improve Khmer architecture and guide young Khmer generations to love and care of all the golden architecture of Khmer.

4. Further Works

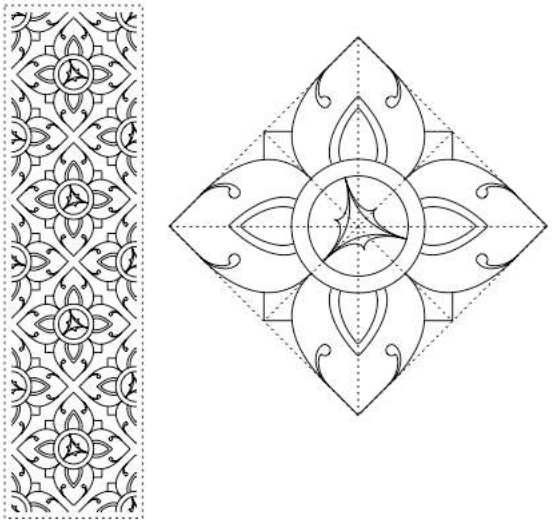


Figure 5. Sample of combination of Khmer ornament using a *Phkar Chan* module

We have many further works to be researched in order to develop our project. You already know that we have done a work of Khmer's flower art. Our goal is not here. We want to have many themes for each Khmer ornament, not just one. Each ornament has properties to change to size and shape of its components according to the art of Khmer. We then move on to combine the related ornaments together depending on any shape architects need. For example, they want the selected ornaments to fit in their provided canvas. Thus, we need to calculate to size of the canvas and how to organize and position each ornament to the canvas. Our goal is to make an easy way that all related fields of Khmer architecture can obtain the ornament to their works.

References

[1] Chan Vitharin and Preap Chanmara, "Kbach: A Study of Khmer Ornament", ISBN: 978-1588860767, Reyum Institute of Arts and Culture, 2005.

[2] Nguonphan Pheakdey, 'Computer Modeling, Analysis and Visualization of Angkor Wat Style Temples in Cambodia'. Ph.D. dissertation. Interdisciplinary Center for Scientific Computing, Heidelberg University, 2009.

Patent Citation Recommendation Based on Topic Model Expansion

Supaporn Tantanasiriwong† **Choochart Haruechaiyasak ‡**

†School of Engineering and Technology
Asian Institute of Technology, Thailand
‡Speech and Audio Technology Laboratory (SPT)
National Science and Technology Development Agency
†Supaporn.tantanasiriwong@ait.ac.th
‡Choochart.haruechaiyasak@nectec.or.th

Abstract

One important task related to patent citation is to recommend suitable citations for a given patent. Patent citation recommendation is difficult due to legal term usage and different skill level in using keyword search among the users. In this paper, we proposed topic model based technique for term expansion to improve the performance of patent citation recommendation. In order to avoid ambiguity from keyword search, we form a query as a document instead of using keywords. Experiment was performed to compare the similarity among Cosine, Jaccard, and KL-Divergence after finishing topic model expansion approach. The framework we proposed is implemented and evaluated its performance using Mean Average Precision (MAP). And the evaluation results show that our proposed approach outperforms the baseline approach.

Keywords: Information Retrieval, Patent Retrieval, Citation Recommendation, Query Expansion, Patent Search

1 Introduction

Patents play a key role in research innovations due to intellectual property protection and for commercial purposes. Patent retrieval is one part of Information retrieval usually aims to retrieve patent documents that satisfy the inventor's needs [9]. It is very difficult to work on patent retrieval due to the different of keywords usage and limited number of terms in query. In this paper, finding patent citation from patent is different from other publication articles since patents typically cite both patents and publications.

Traditionally, we use information retrieval model that is a standard technique for finding the relevant documents by using specific keywords from user's queries [8]. The users become unsatisfied to this since returning the right documents depend on the vocabulary knowledge of user's query. Therefore, we introduce search query by a document instead of using keyword search method from users.

The structure of patent documents contains several parts including title, abstract, claim, description, summary and references which are different from publication papers. Many previous patent search works on developing query patent for automatic query search using parts of patent. Some patents focus on generating query using content from claim [2] but some use the summary [1]. In this paper, we combined title and abstract in order to perform it as a patent query by document. We focus on patent information source as a relationship of patent citation matching.

In context of citation recommendation, various techniques are mentioned in [5] including a discriminative, context-aware, translation model approaches that are used to predict the reference papers from the manuscript. It was proposed by [7] that the ranking for evaluation focuses on the important publications from the number of cited documents. Graph based is one of the most popular techniques applied to use on these papers. However, our paper focuses on recommending bibliography for patent domain which has not been proposed in previous research.

Query expansion has been discussed in [3] and [4]. Both of them used WordNet [10] as a vocabulary based expansion. Although making query expansion by WordNet has not been successful enough, therefore we tried to extend it by topic learning model to solve this problem. Hence, our contribution is to propose a new framework for suggesting bibliography for patent domain based on three main approaches.

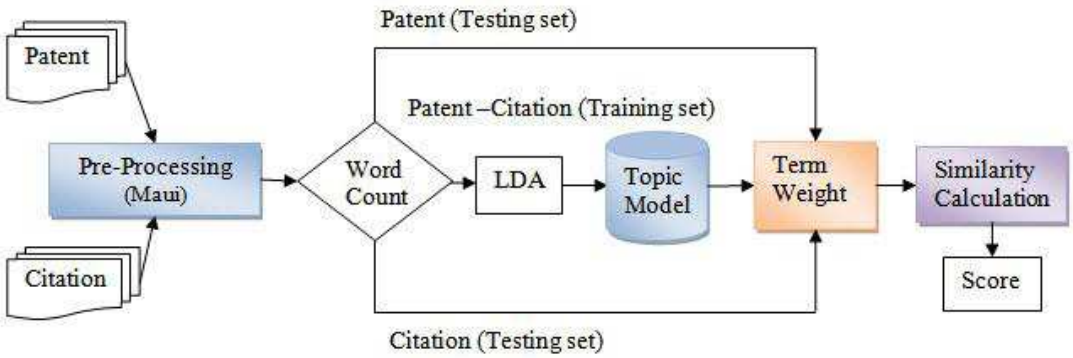


Figure 1. Patent citation matching framework

Firstly, the baseline approach has been implemented as a common mapping technique. Secondly, the manual keyword selection approach where the general words in each paper is removed before performing similarity calculation. Thirdly, the topic model expansion approach is applied using LDA technique to generate topic model in which groups of synonym and hyponym were used to adjust the query weight from the source domain.

The rest of the paper is arranged as follows. Our proposed framework is explained in Section 2. The experiment and results are described with evaluation in Section 3. Finally, the conclusion of the paper is discussed in Section 4.

2 The proposed framework

Our proposed framework consists of three main stages as shown in Figure 1.

In the first stage, we used the key phrase extraction tool called Maui[1] to extract keywords and key phrases through data collection and also removing the stop word. All approaches are implemented using the mentioned pre-processing.

In this second stage, we separate data into two groups, one is training set, and another one is test set. In the framework, we focus on using topic model which is generated from LDA algorithm for training in patent citation documents. Then, patent query testing is transformed to new query using term weight from topic model selection. After that a pair of patent citation test transferred to the similarity calculation for computing the score of citation ranking.

Topic model is a probabilistic model that generated from the LDA algorithm represented as a different topic distribution [7]. Each topic consists of the distribution of words and represented by a probability distribution over terms. Based on the LDA algorithm, it contains the Gibbs sampling methods proposed by [7]. There are three main parameters used in calculation. Such parameters are shown in the probabilistic relationship as follows.

$$P(w_i \mid d) = \sum_{j=1}^{z} P(w_i \mid z_i = j) P(z_i = j \mid d) \qquad (1)$$

where $P(w_i \mid d)$ is the probability of arbitrary word w_i given by a document d . z_i is represented as a latent topic over word distribution into a given document.

In recommending patent citation system, a set of document in collection is given and denoted by $C = \{d_0, d_1, ..., d_{i-1}\}$ which each document consists of a list of words denoted by $d_i = \{w_0^i, w_1^i, ..., w_{|d_i|-1}^i\}$ and $|d_i|$ is the total number of words in d_i.

In the topic model expansion, topic model training denoted by $T = \{T_1, T_2, ..., T_j\}$ is a set of topics where T_j represents an arbitrary topic. The topic contains a probability of words denoted by $T_j = \{w_0^j, w_1^j, ..., w_{m-1}^j\}$ where w_{m-1}^j is a probability of word distribution in each topic. The

[1] http://code.google.com/p/maui-indexer/

appropriate topic model needs to be selected beforehand. After that, we use a set of words input query combine with a list of words derived from topic model selection. Then, the new patent query can be transformed by $d_i^{/} = \{w_0^i, w_1^i, \ldots, w_{|d_i|-1}^i\}\{w_1^j, w_2^j, \ldots, w_{m-1}^j\}$ as a query expansion using the topic model expansion.

In the third stage, we presented Cosine, Jaccard, and Kullback-Leibler (KL) to measure the similarity distance among two documents. Given d_i = { $w_i^0, w_i^1, \ldots, w_i^{T-1}$ } and d_j = { $w_j^0, w_j^1, \ldots, w_j^{T-1}$ }, where d_i and d_j are represented as a set of word w in document d_i and d_j. Every similarity matching calculated as a vector and its co-occurrence distribution. The number of similarity was ranked after using the measures defined in Equation (2), (3), (4).

Cosine Similarity:

$$Sim(d_i, d_j) = \frac{d_i * d_j}{|d_i| * |d_j|} = \frac{\sum w_i * w_j}{\sqrt{\sum w_i^2}\sqrt{\sum w_j^2}} \quad (2)$$

Jaccard Co-efficient:

$$Sim(d_i, d_j) = \frac{d_i * d_j}{|d_i|^2 + |d_j|^2 - d_i * d_j} \quad (3)$$

Kullback-Leibler (KL) Divergence:

$$KL(d_i, d_j) = \sum d_i(x) \log \frac{d_i(x)}{d_j(x)} \quad (4)$$

3 Experiments and results

3.1 Data collections

We collected data from Google Patent Search[2] which includes patent documents from the USPTO (US Patent Office) [6]. The database contains 981,948 patents published articles from 1993 to 2000. In our experiments, we randomly selected 200 patent citations for each of the following ten innovative technologies: OLED, Carbon Nanotubes, Fuel cell, GIS, RFID, Alternative energy, 3D Prototype, PDA, Solid state battery, Brain interface. In total, the dataset consists of 2,000 patents with 23,018 citations. In addition, we also separate a pair of patent citation into 100 patents for a test set and leave others 1,900 for training model.

3.2 Result and evaluation

There are three main approaches presented in this paper as shown in Table 1, 2, 3 measured by three performance metrics which are precision, recall, and F1-measure.

Baseline approach: The baseline approach is implemented as a traditional document mapping technique using the frequency of words to represent as a vector from a document. The total number of unique words obtained from the baseline is 14,848 patent-citations.

Keyword selection approach: This approach focuses on filtering with the common word normally used in all document (e.g. provide, determine) then only main keywords or key phrases will be utilized to perform document mapping calculation. The total number of unique words that were used in this approach after filtering is equal to 14,513.

Topic model expansion approach: We use the linguistic analysis tool LingPipe[3] to generate a topic model from training data set. Then, the term weight expansion is extended to a set of term query in test set. Finally, the patent query is regarded as a new feature expansion by topic model. After running topic model with various ranks, i.e. 50, 100, 150 and 200, we found that the topic at 200 acquired the highest value in the local maximum point.

Table 1. Baseline mapping

Similarity Rank@50	Mean Precision	Mean Recall	Mean F-Measure
Cosine	0.077	0.486	0.119
Jaccard	0.080	0.468	0.122
KL	0.078	0.487	0.120

[2] http://patents.google.com

[3] http://alias-i.com/lingpipe/

Table 2. Manual Keyword Selection

Similarity Rank@50	Mean Precision	Mean Recall	Mean F-Measure
Cosine	0.078	0.493	0.120
Jaccard	0.082	0.479	0.126
KL	0.080	0.486	0.123

Table 3. Topic Model Expansion @Topic200

Similarity Rank@50	Mean Precision	Mean Recall	Mean F-Measure
Cosine	0.081	0.508	0.124
Jaccard	0.087	0.506	0.131
KL	0.082	0.506	0.126

In overall consideration, Jaccard co-efficient method excels compared to other similarity methods. We found that the topic model expansion showed better performance than both manual keyword selection and baseline approach.

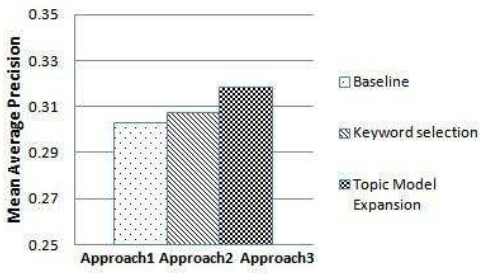


Figure 2. Performance using Jaccard co-efficient

In Figure 2, we focus on Mean Average Precision (MAP) obtained from Jaccard matching in different approaches where MAP is considered the rank position of a relevant document. The equation is as follows.

$$MAP = \frac{\sum_{q=1}^{Q} AveP(q)}{Q} \qquad (5)$$

Where $AveP(q)$ is average precision in each query and Q is the number of queries.

The final result indicates that a topic model expansion approach could help increase the performance of patent citation matching over the baseline approach by 1.5% and manual keyword selection approach by 1.1%.

4 Conclusions

In this paper, we propose patent citation recommendation framework by processing patent citation information and implementing the topic model as a semantic word expansion method. In this case, the methods of similarity matching such as Cosine, Jaccard, and KL-divergence are used to calculate and compare the performance of three main approaches. It can be concluded that the topic model expansion approach could help improve the accuracy of patent citation domain matching. Although the result is less than what we expected, however in the future we are planning to use other feature extraction in order to improve patent matching with the larger number of data collection.

References

[1] X. Xue and W.B. Croft, Automatic query generation for patent search. Proc. of the CIKM'09 conf., pp. 2037-2040, 2009.
[2] X. Xue and W.B. Croft, Transforming patents into prior-art queries. Proc. of the SIGIR'09 conf., pp. 808-809, 2009.
[3] J. Zhang, B. Deng, and X. Li, Concept Based Query expansion using WordNet. Proc. of the AST'09 conf., pp.52-55, 2009.
[4] W. Magdy and G.J.F. Jones, A study on query expansion methods for patent retrieval. Proc. of the PaIR'11 conf., pp. 19-24, 2011.
[5] Q. He, J. Pei, D. Kifer, P. Mitra, and C. L. Giles, Context-aware citation recommendation. Proc. of the 2010 WWW conf., pp. 421- 430, 2010.
[6] United States Patent and Trademark Office. http://www.uspto.gov/.
[7] T. Griffiths, Handbook of latent semantic analysis. Probabilistic topic models, Lawrence Erbaum Associates, 2007.
[8] C. J. Van Rijsbergen, Publication Book. Information Retrieval. Butterworth-Heinemann Newton, ACM, 1979.
[9] A. Fujii, M. Iwayama,and N. Kando, Overview of the Patent Retrieval Task at the NTCIR-6 Workshop. Proc. Of NTCIR-6 Workshop Meeting conf., pp. 359-365, 2009.
[10] G. Varelas, E. Voutsakis, P. Raftopoulou, E.G.M. Petrakis, and E. Milios, Semantic Similarity Methods in WordNet and their Application to Information Retrieval on the Web. Proc. of the WIDM'05 conf., pp.10-16, 2005.

Chinese and Japanese Word Learning System by Estimation of Word Difficulty

Peiqing Zang† Hiromitsu Shiina‡ Nobuyuki Kobayashi* Fumio Kitagawa‡

†Graduate School of Informatics
Okayama University of Science, China
‡Faculty of Informatics, Okayama University of Science, Japan
* Faculty of Human Science, Sanyo Gakuen University, Japan
zangpq@gmail.com

Abstract

Many foreign students are also enrolling universities in Japan. When foreign students undertake lectures in Japanese, there is an assumption that they understand a good deal of Japanese. However, they do not understand all of the vocabulary needed for lectures. We believe that a Japanese word preparation learning system for providing the vocabulary required for a lecture is required. In addition, there is a change toward using mobile devices for learning, such as iPhone, iPad, Android Tablets, and son. As such, we need to provide a simple method for learning. This research is developing a system for practicing Japanese and Chinese vocabulary using terms taken from a text for beginners, as well as technical term taken from VOD lectures on a database server that have open access to documents such as lecture videos and slides. We are also utilizing the difficulty level of Chinese and Japanese terms to determine the order of appearance and to extract incorrect terms in multiple choice vocabulary problems. To determine the difficulty level of each word, we used support vector machine (SVM) to estimate the difficulty of words with no preexisting difficulty level based on words whose level is already established, and then used a bootstrap method which made use of the estimated difficulty levels of these words.

Keywords: Word difficulty, Machine Learning, Support Vector Machine, Learning Support System

1 Introduction

Many foreign students are now enrolling universities in Japan. It is assumed that a foreign student attending lectures in Japanese will be quite advanced in the study of the Japanese language, but in reality students often arrive after a short period of Japanese language study and have an incomplete grasp of the language. Because of this, they cannot understand all of the words used in their lectures. We believe these students need a self-learning system, before and after class, the Japanese terms used in their lectures.

Meanwhile, mobile devices such as the iPhone, iPad, and Android Tablet are changing the learning environment, and because of this we believe that simple learning methods are desirable. Many software programs have already been developed to help learners study foreign language vocabulary. Among these, nearly all of the multiple choice quiz applications are structured to present a target language prompt with answer choices in the user's native language. Few applications present questions with a native-language prompt and target-language answer choices, and even fewer combine the two methods.

In this research, we address this by creating a system structure that combines Japanese language and Chinese language vocabulary training systems. The goal is to improve the results of vocabulary training, in particular by using Japanese-to-Chinese study sessions and multiple-choice questions in combination with Chinese-to-Japanese multiple choice questions and answers.

Our vocabulary training system was developed for the Android Tablet device, and in order to create this system it was necessary to estimate the relative difficulty levels of the words used as vocabulary items and incorrect answer choices.

To classify words by difficulty, we used the difficulty levels from the Chinese Proficiency Test (HSK[1]) for Chinese and the Japa-

nese-Language Proficiency Test (JLPT[2]) for Japanese. To determine the difficulty level of each word, we use SVM[3] to estimate the difficulty of words with no preexisting difficulty level based on words whose level is already established, and then used a bootstrap[4] method which made use of the estimated difficulty levels of these words.

Finally, for the subjects to be studied, we chose categories of words which foreign students would need to start life in Japan as well as those which would be useful to learn before attending lectures at a Japanese university.

2 Estimating vocabulary difficulty levels

2.1 Estimating vocabulary difficulty levels

When using data from the dictionary to estimate the level of difficulty for vocabulary, we theorized that the words used in the dictionary definition will be either easier than, or the same difficulty as, the entry word. Under this theory, the difficulty level distribution seen in the words that appear in the definition correlates to the level of difficulty of the entry word. We then carried out machine learning with an SVM, using the difficulty level distribution of the words in the definition as parameters.

This allowed us to estimate the level of difficulty of all the dictionary entry words. The difficulty estimation learning algorithm is a bootstrap method which applies words for which difficulty has been determined from teacher data to the definition, and then repeatedly assesses the difficulty of words for which difficulty has not been determined. The relationship between teacher data and the dictionary and the SVM learning process is shown in Figure 1. The following shows the steps involved:

(Step1)SVM learning with the initial learning data

If the level of difficulty of the words in the dictionary definition had not been estimated, the level of difficulty of those words was exempted from the difficulty level distribution, then learning parameters were made and learning was carried out by an SVM.

When the level of difficulty of the words was classified as L, and the frequency for each word difficulty level appearing in the semantic

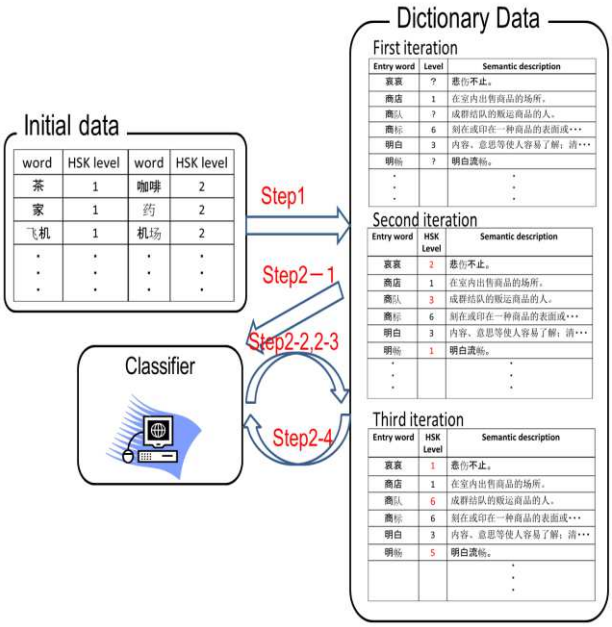


Figure 1. A Summary of the Difficulty Learning Process

description for entry word w was D(l, w), l = 1, ..., L, then the standardized word difficulty frequency distribution was defined as follows:

$$DR(l,w) = \frac{D(l,w)}{\sum_{i=1,\ldots,L} D(i,w)}.$$

The SVM learning parameter (level of word difficulty of (w, DR(1, w), DR(2, w), ..., DR(L, w)) was composed of the level of difficulty the word w and L in the difficulty level distribution in the dictionary definition.

(Step2)Repeated SVM learning

(Step2-1) If the level of difficulty of the entry word was not in the initial data, we estimated it using the difficulty level distribution of the words in the dictionary definition.

(Step2-2) If the levels of difficulty for all of the words in the definition had not yet been determined, we made a difficulty level distribution from words whose difficulty level was known, then we estimated the difficulty level of the entry word.

(Step2-3) We updated the difficulty level of the words in the definition, then carried out classification learning with an SVM using parameters combining the level of difficulty of the entry words of the initial learning data.

(Step2-4) We then repeated (Step2-3). If the changes in results for the estimated level of difficulty converged, then we were finished.

2.2 Determining difficulty levels in Chinese

We used the test levels from the HSK as difficulty levels for Chinese. For the initial dataset[5],[6], we used a difficulty level list compiled from lists of words which have appeared on the test in the past. In addition, we used Hanyu Da Cidian Dictionary as Chinese dictionary[7]. Difficulty levels of words not on this list were determined by extrapolation. The difficulty levels of Chinese terms were determined by the same method discussed for Japanese below.

The Chinese method differs from the Japanese in the way dictionary entries are partitioned into individual words when difficulty levels are assigned; for this purpose, we have used a form of morphological analysis[8] suited to Chinese.

The HSK is divided into six levels, with Level 1 being the easiest and Level 6 the most difficult. As an example, Figure 2 shows how the difficulty level is determined for「加油」(jiayou). In Figure 2, the zeroth difficulty level estimation assigns difficulty levels only to the words whose difficulty is already known, and the SVM estimates a difficulty level of 4 based on the difficulty level distribution (4/11, 1/11, 1/11, 1/11, 2/11, 2/11). The first estimation uses the zeroth estimation to estimate the difficulty of the dictionary entry word and then uses that information to assign a difficulty level to each word; here, the SVM estimates a difficulty level of 5 based on the difficulty level distribution (4/17, 4/17, 2/17, 3/17, 2/17, 2/17). Note that symbols and words whose difficulty levels have not been determined are labeled “0” in Figure 2.

The word 「汽车」(feiji) appears in the dictionary definition used in the difficulty level estimations but is not included in the teaching data, and so it is assigned Level 2 during the first difficulty level estimation; it is then used in the estimation of 「加油」.

3 Selection of vocabulary items

3.1 Selection of vocabulary categories

Vocabulary categories were selected from the New Standard Japanese for Sino-Japan Communication Primary[9] and arranged in order of which were considered the most crucial to life in Japan. Lessons are divided into tenth study categories: (1) greetings; (2) transportation; (3) food; (4) people; (5) numbers; (6) facilities;

Figure 2. Estimation Process of Difficulty for「加油」 (jiayou) in Chinese

Chinese vocabulary						Japanese vocabulary					
Category	Problem Word	HSK Level	Correct Word	Incorrect Word1	Incorrect Word2	Category	Problem Word	JLPT Level	Correct Word	Incorrect Word1	Incorrect Word2
Greetings 问候	你好 Hello	1	こんにちは/4	おはよう/4	こんばんは/4	Greetings 挨拶	こんにちは	4	你好/1	早上好/1	晚上好/1
	再见 Goodbye	1	さようなら/4	ひさしぶり/4	また明日/4		おはよう	4	早上好/1	晚上好/1	你好/1
	早上 Good morning	1	おはよう/4	こんばんは/4	こんにちは/4		さようなら	4	再见/1	明天见/1	好久不见了/1
	明天见 See you Tomorrow	1	また明日/4	お世話になりました/3	ひさしぶり/4		また明日	4	明天见/1	好久不见了/1	不用谢/1
Transportation 交通工具	火车 Bus	1	列車/2	トラック/2	地下鉄/4	Transportation 交通機関	地下鉄	4	地铁/3	无轨电车/2	直达快车/2
	公共汽车 Bus	2	バス/1	地下鉄/4	タクシー/2		列車	2	火车/1	无轨电车/2	特别快车/2
	地铁 Subway	3	地下鉄/4	列車/2	トラック/2		タクシー	2	出租汽车/6	地下鉄/4	バス/1
	出租汽车 Taxi	6	タクシー/2	バス/1	列車/2		バス	1	公共汽车/2	火车/1	地铁/3
Food 食物	曲奇 Cookies	2	クッキー/4	パン/4	ビスケット/1	Food 食べ物	パン	4	面包/3	饼干/4	蛋糕/4
	面包 Bread	3	パン/4	ビスケット/1	クッキー/4		クッキー	4	曲奇/2	蛋糕/4	蜂蜜/2
	蛋糕 Cake	3	ケーキ/2	クッキー/2	ゆで卵/1		ケーキ	2	蛋糕/3	曲奇/2	面包/3
	饼干 Biscuit	4	ビスケット/1	パン/4	ベーコン/2		ビスケット	1	饼干/4	面包/3	蛋糕/4

Figure 3. Examples of Word lists for Chinese and Japanese vocabulary with incorrect words

(7) the week; (8) climate; (9) occupations; and (10) clothing.

3.2 Selection of vocabulary items and incorrect answer choices

Words which appear as questions were selected from the categories discussed above, and their difficulty levels in each language were used to arrange the words in order of increasing HSK difficulty in the case of Chinese and increasing JLPT difficulty in the case of Japanese. We think that when the difficulty level of both Japanese terms and their Chinese translations is low, they are conceptually easier to understand, and that these should be the easiest terms to study. Even if the Japanese is easy, it is still difficult to understand if the Chinese is difficult, and we suppose

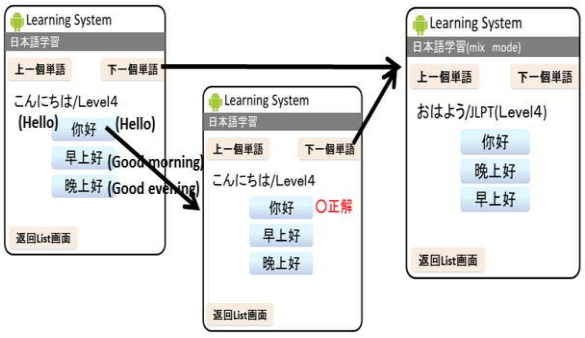


Figure 4. Screen transition of the normal mode

that studying is most difficult when both terms are difficult.

Incorrect answer choices were chosen based on the following criteria:

(1) Incorrect choices are chosen from the same category as the answer.

(2) Incorrect choices are of a similar level to the answer.

However, during the actual development of the system, incorrect answers were first selected automatically by the computer and then checked manually before being included. This is because the computer does not consider the meaning of the words(Figure 3).

4 The vocabulary training system

The vocabulary training system offers multiple-choice vocabulary questions with Chinese prompts and Japanese answer choices, as well as questions with Japanese prompts and Chinese answer choices. The following two training methods are provided:

(1) Normal Mode: The user studies multiple choice questions from one language at a time. If the user is studying Chinese, he or she will move on to the next Chinese prompt (Figure 4).

(2) Mixed mode: Correct answers appear again as questions. If the user is studying in Chinese, there will be a question on the Japanese word he or she selected, and the program will return to a Chinese question prompt when it is completed (Figure 5).

5 Future Works

In this research, we determined the difficulty of Chinese and Japanese vocabulary terms via an SVM and developed a training system which utilized these difficulty levels in creating incorrect answer choices. Study terms included words selected from the New Standard Japanese

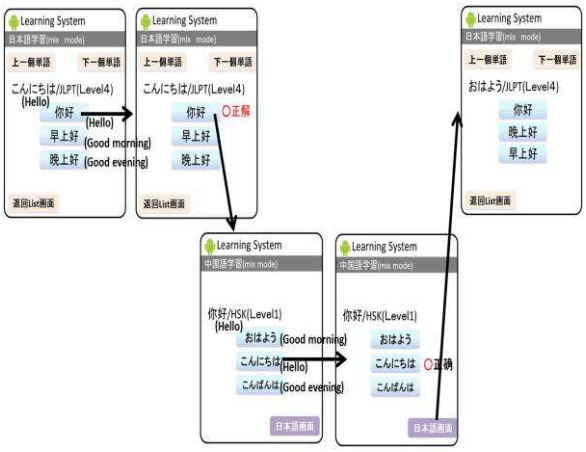


Figure 5. Screen transition of the mixed mode

for Sino-Japan Communication Primary[15] and specialized terms used in university lectures. Furthermore, the vocabulary training method we developed not only includes a mode where the user studies a series of either Chinese or Japanese prompts, but also a mode in which some prompts appear in the opposite language. We consider the combined use of these two methods will improve results. In the future, we hope to construct a vocabulary training system which fits each student's proficiency level. We also plan to work toward developing a training system which tells students how their vocabulary studies are progressing in real-time as they use it.

References

[1] http://www.hsk.org.cn/Default.aspx

[2] Japanese-Language Proficiency Test, http://www.jlpt.jp

[3] V. Vapnik, Statistical Learning Theory, Springer, 1998.

[4] Bradley Efron, Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics 7 (1), pp 1-26, 1979.

[5]http://ysnhrk.blogspot.jp/2012/12/chinese-word-list-for-hsk.html

[6]http://chugokugo.shisyou.com/side_menu1_10.html

[7]Hanyu Da Cidian, Publishing House of the Chinese Dictionary, 1998.

[8] Natural Language Processing & Information Retrieval Sharing Platform,http://www.nlpir.org/

[9] New Standard Japanese for Sino-Japan Communication Primary, Peoples Education Press, 2005.

A web based programming learning management system with JavaScript libraries

Shimpei Matsumoto† **Shuichi Yamagishi**† **Tomoko Kashima**‡
†Faculty of Applied Information Science, Hiroshima Institute of Technology
‡Faculty of Engineering, Kinki University
†{s.matsumoto.gk, s.yamagishi.if}@cc.it-hiroshima.ac.jp
‡kashima@hiro.kindai.ac.jp

Abstract

A web based programming learning management system is now planned to be developed aiming for the holistic improvement of understanding for programming. The system is for students who are not good at programming and are with insufficient understanding, and the system aims to train their skills by the after class hour learning. Before the authors develop the software, this paper addresses to give the part of the specification of the software. Especially as the first step, this paper examines what kind of students are no good at programming, and what needs to improve the understanding of the students. The authors consider that the trends and characteristics of students with low programming skill need to be shown in supporting their learning. The short exams were conducted in several times in lectures related to programming for the first and second grade students currently being held, and the results were examined depending on the achievement degree of students. From the analysis, the valuable speculation about students who are no good at programming was obtained.

Keywords: Programming, Algorithm, Learning support, Web, JavaScript, Puzzle, Entertainment

1 Introduction

Information systems technology is essential for today's highly-sophisticated information society, and the demand in society for engineers of IT has been extremely high. Therefore various educations on IT have been practiced in the departments of informatics or related fields in higher educational institutions mainly universities. Specifically in the efforts of these institutions, educations in programming related subjects including software engineering, numerical computation, algorithm theory, and object oriented modeling has been emphasized because these programming techniques is absolutely necessary to build information systems. Many companies require high-level programming skill for university students majoring informatics. Therefore each educational institution should devise an efficient educational method more for programming related lectures.

The authors have taught some programming related subjects, introduction to algorithm for freshmen, introduction to the C programming language for freshmen, and applied algorithm for sophomores with traditional common lecture style for several years, but the authors could not have provided enough guidance to overcome a lack of the understanding for students who are not good at programming. As far as the authors verify, every year there are inevitably students in definite proportions who are not good at programming, and the trend is thought to be similar between each grade. Interestingly, the reference [2], a quite famous paper on the web in spite of it was not officially published, reported a similar result; students not good at programming certainly exist at a constant rate regardless of gender, age, and educational level.

Based on the above background, the authors began educational research of programming related subjects for improving the overall level of programming technique. Concretely speaking, construction of programming self-studying environment with visual and intuitive operation, which can generate interest in programming, is targeted as the final goal. As the self-studying environment, an Internet accessible programming learning management system (PLMS) is planned to be developed and introduced, which enables to enjoy learning of the basic perspective of programming by giving entertainment factor. Firstly this paper clarified the character and functions of PLMS to implement by objectively

analyzing all students' programming skill with some examinations conducted in actual lecture. From the result of the analysis, beneficial knowledge to design a PLMS was disclosed. With the knowledge, prototype of a PLMS was developed on the basis of user friendly learning function like puzzle games. The PLMS focuses on students who want to be good at programming but right now they are not good at. This paper presents game like programming learning questions with Google Blockly and enchant.js, JavaScript libraries, which aims to automatically move a robot from start point to goal point by giving an algorithm consisted of some pieces with proper order, each of which corresponds to an instruction to move.

2 Related Works

Earlier researches on programming are roughly classified into two subjects: the one is on the teaching method for programming learning, and the other is on the development of programming learning support system. For example, Fukamachi showed the typical mistakes of which liberal arts students are likely to misunderstand for learning computer programming by giving specific examples, and addressed the improvement of lecture for liberal arts students [1]. Fukamachi indicated that the major factors for preventing programming learning are lack of ability to interpret mathematical formulae, and lack of experience to express one's thinking with mathematical formulae and training to do this. Fukamachi also said it is the critical problem that students cannot solve errors themselves at the stage of execution or compilation.

As it seems that programming strongly depends on one's natural inclinations, Dehnadi et al. insisted that judgment whether a student is good at programming or not is possible in the early stage, before the training of programming, by using relatively simple test [2]DDehnadi et al. said that the percentage of people with the talent for programming is universal not depending on gender, age, and educational level. Dehnadi et al. also concluded that disposition, not to seek meaning for conclusion as long as thinking is consecutive, is important for programming, and the rate of people who have a talent to enjoy the training of programming is approximately 40 percent.

There are many practices achieving remarkable results. As one of the practices, Tada et al. conducted a research on lecture method which was introduced repetitive study on programming education [3]. This method can effectively memorize frequent patterns in programming for students. This paper considers that entertainment factor is important for education to continue repetitive study. As a practical example of education focusing on the entertainment factor, Komatsu provided a programming education of the basic concept on object-orientation with ActionScript [4]. He showed the effectiveness through the practice of game creation, and notably concluded that the characteristic of ActionScript contributed to an intuitive understanding of programming. Considering these efforts, this paper develops an Internet accessible PLMS which provides an intuitive programming learning function basing on the entertainment factor, and social web functions like SNS.

3 Our Current Work

Right now the authors have taught some programming related subjects for freshmen and sophomores. In the lecture held at first semester of freshman, the basic idea of algorithm, such as description of flowchart, trace of variable value, and the C programming language (from programming syntax to the basic of the array) have been mainly introduced. Subsequently in the lecture held at second semester of freshmen, students have been given the overall idea to interpret and to write the C programming language. Then in the lecture to practice the applied idea of algorithm held at first semester of sophomores, some famous algorithms to seek and to operate data, such as binary search and quick sort, and programming technique to implement these algorithms, have been introduced.

Among these lectures, the first lecture for the introduction of the algorithm is carefully planned to reduce the feelings of resistance toward programming. Concretely speaking, a web based learning game named "Algologic" is used [5]. Algologic is a Web-based software program to learn logical thinking, the fundamental skill to program, with it just like playing games. As a result of using Algologic, a lot of positive feedback was received from questionnaire. However in spite of utilizing Algologic, students who could not catch up with the progression of lecture continued to increase as the lecture went

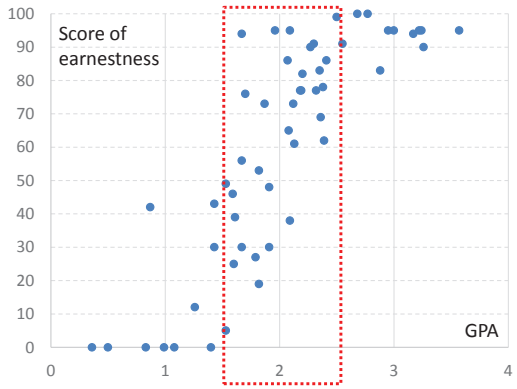


Figure 1. Relation between GPA and Score of Earnestness

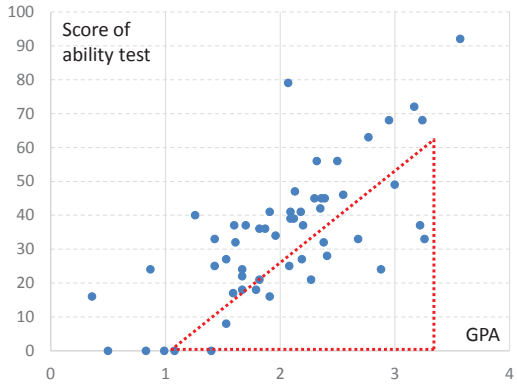


Figure 2. Relation between GPA and Score of ability test

on. The achievement test conducted on the final stage showed the existence of many students who could not understand the basic idea of programming at all.

4 Preliminary Investigation and Results

Several tests were conducted in our lecture, and each one's programming skill was objectively evaluated. Also the programming skill and the relation between other factors were examined. Based on the analysis results of the examination, students targeted by our PLMS and functions which should be implemented to the PLMS were clarified.

The lecture of algorithm for sophomores was intended for analysis. The main reason this paper chose the lecture was that the characteristics of students would be relatively noticeable by the previous learning. The other reason this paper chose the lecture was that students targeted by this analysis did not take our lectures, and they were taught by traditional lecture method. In this lecture, there were 55 students. The data used by the analysis is shown below. At first the results of examinations, 12 mini examinations and 2 achievement tests were used, and each test evaluated on a 100-point scale. The questions of each mini examination are based on the content of the previous lectures, so student can solve the exercise as long as he/she fully and correctly understood. On the other hand, the achievement tests include some practical exercises, so students should understand the essence of algorithm and programming. Second, the data about evaluating efforts for the lecture were used: attendance records and submission records of assignments. 12 assignments required to submit, and the submission rate of the assignments were used by the analysis. Additionally, each student's grade point average was referred.

GPA includes a degree of understanding and earnestness for all subjects. To make the value of GPA easy to understand, an index to evaluate earnestness was defined, and it is named as "Score of earnestness". The score of earnestness was evaluated on a 100-point scale, and was calculated by the average of attendance rate and the submission rate of assignments. The relation between GPA and Score of earnestness is shown in Figure 1. From Figure 1, the cause of GPA under 1.5 is thought to be a lack of earnestness because most students' score of earnestness with GPA below 1.5 is low. The correlation between GPA and Score of earnestness seems strong from 1.5 to 2.5 of GPA. Most students with GPA over 2.5 have high value of Score of earnestness, so earnestness is thought to be a necessary condition to obtain GPA over 2.5.

Next second index was defined to evaluate programming achievement degree, and it is named as "Score of ability test". The index was calculated by the average of all tests: 12 mini examinations and 2 achievement tests. The relation between the index and GPA is shown in Figure 2. From Figure 2, we can see relatively strong correlation. The noteworthy point from the result is that there were some students lacking sufficient programming knowledge even they had GPA over 2.5. They are thought not to be interested in programming but earnest, so the authors consider that they will self-study the learn-

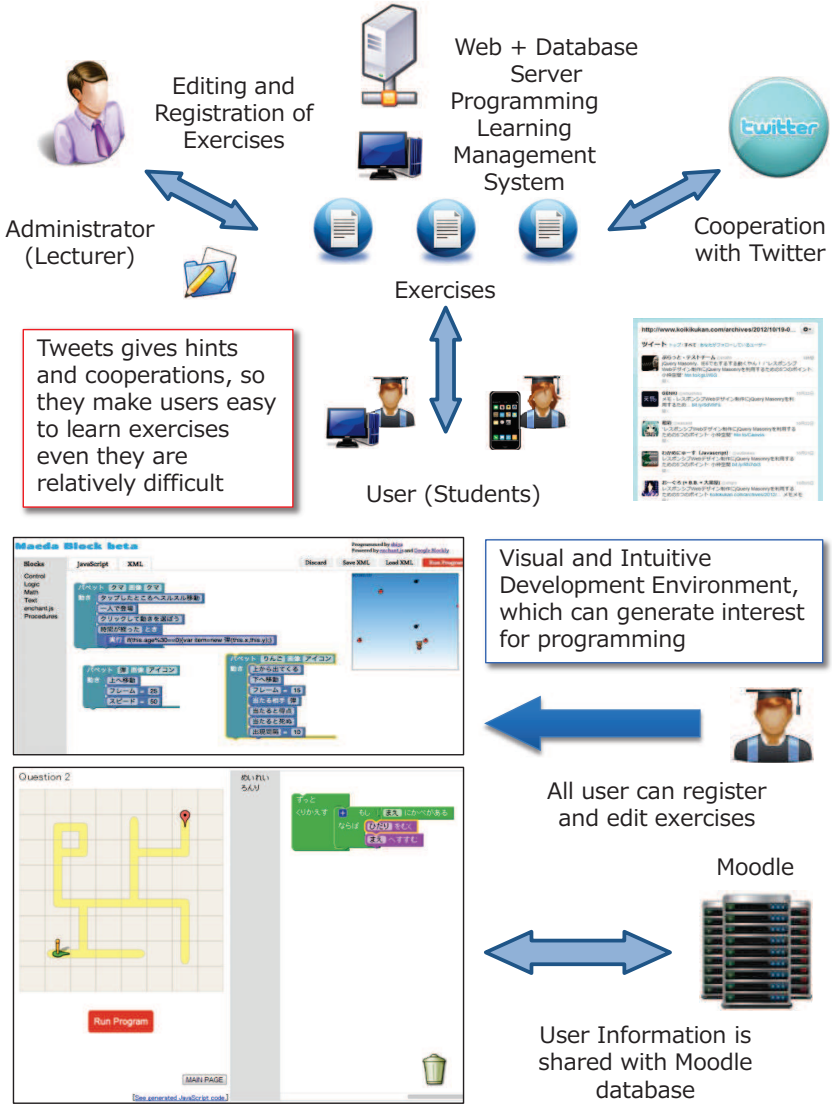


Figure 3. Proposed System

ing of programming if they can take an interest because they are earnest.

5 System Development

5.1 Overview of the Proposed System and its Implementation

This paper focuses on students who can self-study but are not interested in programming at this time. A PLMS is developed which enables users to learn programming like game by considering the characteristics of Algologic and by using the functions of Google Blockly [6] and enchant.js [7]. The overview of the proposed system is shown in Figure 3. As above mentioned, Algologic is a puzzle game to learn logical thinking available on the web, and the user aims to automatically lead a robot to the goal point by giving action rule which is constructed with some pieces, each of which corresponds to a certain rule such as moving forward, turn, repetition and branching. Google Blockly is a JavaScript library which can generate JavaScript code visually by assembling some parts corresponding to rules, and the user interface is similar with Algologic. enchant.js is a framework for developing simple games and applications in HTML5 and JavaScript. By combining Google Blockly and enchant.js, an open source library named Maeda Block [8] was developed. This paper uses Maeda Block and Google Blockly Maze to implement visual programming envi-

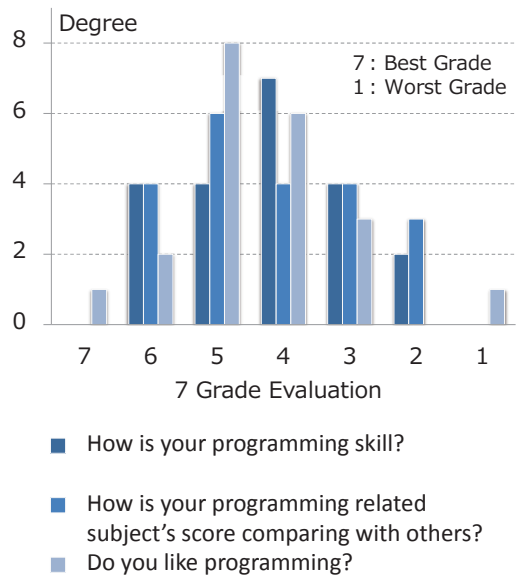


Figure 4. Self-Evaluation about Programming

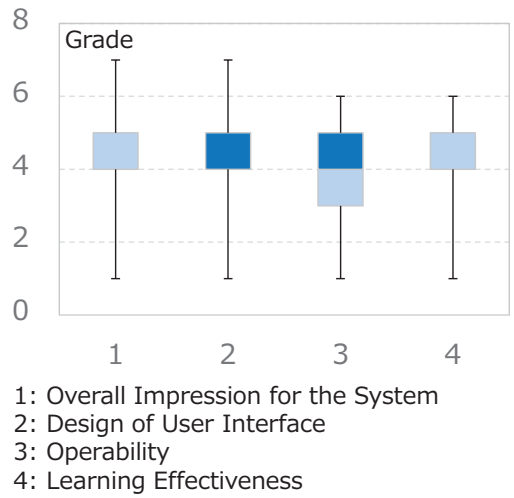


Figure 5. Evaluation for Whole System

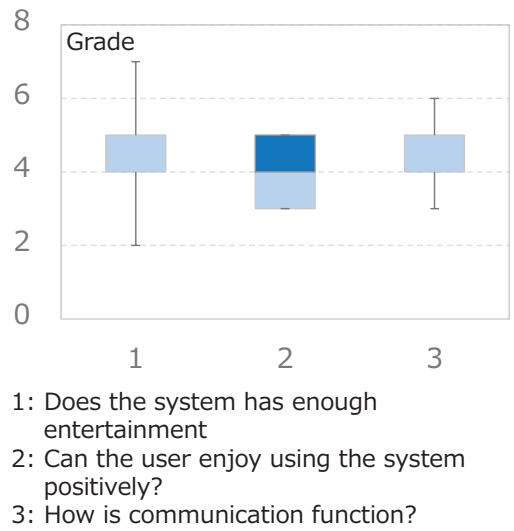


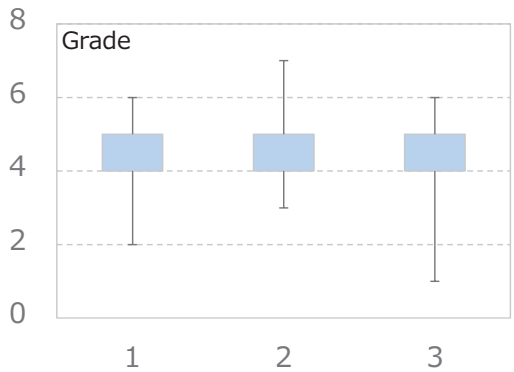
Figure 6. Evaluation for Entertainment

ronment, and provides puzzle based programming questions similar with Algologic. Additionally the proposed system implements a bidirectional communication function like SNS as a new feature. The proposed system runs on a Linux server with Apache 2.2.20, MySQL 5.1.63 and PHP 5.3.6-13. Each user can create, edit, and register the exercise, and challenges an exercise created by another user, and each user can freely post his/her comment for another user and exercises. The exercises are uniquely managed by a database newly developed. With the function of enchant.js, the proposed system is available on Smartphone, and realizes virtual communication between each user through Twitter.

5.2 Evaluation

Evaluation experiment of the proposed system was performed with 16 university and graduate students in informatics as subjects. Before the experiment, attitudes toward programming were firstly surveyed for subjects. The result is summarized in Figure 4 which shows the distribution of subjects for 3 typical questions according to grade. From the distributions of the scores for questions, "programming related subjects" and "programming skill", the subjects are thought to be a group with intermediate programming skill. With the group of subjects, evaluation results for the proposed system (see Figure 5), for entertainment (see Figure 6), and for the possibility to use introductory education (see Figure 7) were obtained. In each box plot as shown in Figure 5 - 7, the range from 25 percentile to 50 percentile is filled with light blue, and the range from 50 percentile to 75 percentile is filled with dark blue.

Figure 5 shows the evaluation for whole system. The evaluation was diverse because the proposed system was a prototype, but it was found that the user interface needs to be improved. In the question relevant to user interface, it was found that some measures and manuals about Google Blockly need to be prepared because the operation takes getting used to. Figure 6 shows that the entertainment oriented design is relatively positive, but the evidence that students will



1: Will the user be interested in programming ?
2: Will the user understand the logical thinking by using this?
3: Will high school, junior high school students be interested in information engineering by using this?

Figure 7. Evaluation for Possibility to use Introductory Education

positively use the proposed system was not so strong. Some of the subjects said only the form of exercise will be tired. To meet the issue, there were some request to register various types of exercises, for example, competition game and battle game with a monster by using an algorithm. About the function of SNS, there were many feedbacks that the user might not be able to positively learn unless the exercises have a kind of game element such as collection and trade function of some kinds of items (accessories). Figure 6 shows negative feedback for programming training, but positive for logical thinking. A subject commented that the best way to train programming skill is to write a lot of program actually.

6 Conclusion

This paper addressed to support self-study of programming. Firstly to grasp students' programming skill objectively, the relation between achievement score and GPA was examined by conducting mini tests several times in the lecture of algorithm for sophomores. In relation to this, an index to evaluate earnestness was defined, and the relation between the index and GPA was examined. As the result, it was found that sufficient attendance and submission of assignments are necessary conditions. This paper confirmed the group of students lacking programming skill but earnest, and PLMS was developed for the group. Considering Algologic, the proposed system which can provide exercises of algorithm like puzzle game was designed. From preliminary experimental result, there were many positive feedbacks, but there were some points to be improved. Students who have enough programming skill tended to be negative, on the other hand, students who are not good at programming tended to be positive. The questionnaire result showed entertainment based visual programming environment with social communication capability was efficient, and the target of the proposed system was proper.

Acknowledgment

This work was supported by Japan Society for the Promotion of Science, Grant-in-Aid for Young Scientists (B), No.13304922.

References

[1] Syuichi Fukamachi, A study of computer programming education for arts students, Annals of fukuoka international university, No.23, pp.39-45 (2010), In Japanese.

[2] Saeed Dehnadi and Richard Bornat, The camel has two humps (working title), http://www.eis.mdx.ac.uk/research/PhDArea/saeed/paper1.pdf, 2012/06/25 referred.

[3] Tomomasa Tada and Hiroyuki Maruta, The course design with repetitive lessons for education of programming, Bulletin of Kyoto University of Education, No.116, pp.123-134 (2010), In Japanese.

[4] Koji Komatsu, The education of programming by Action Script, The Review of Business Administration, Vol.17, No.1, pp.211-226 (2007), In Japanese.

[5] Japan Electronics and Information Technology Industries Association, Algologic – an algorithm experience game, http://home.jeita.or.jp/is/highschool/algo/, In Japanese, 2012/06/25 referred.

[6] Google Blockly, https://code.google.com/p/blockly/, 2013/06/13 referred.

[7] enchant.js, http://enchantjs.com/, 2013/06/13 referred.

[8] Maeda Block, http://junk.wise9.jp/maeda/, 2013/06/13 referred.

Estimating Degree of Concentration through Activity Recognition: Use in a Classroom with SNS

Tomoaki UEDA† **Hiroshi SUGIMURA‡**

Tetsuo TANAKA† **Kazunori MATSUMOTO†**

†Department of Information and Computer Sciences, Kanagawa Institute of Technology
‡Department of Home Electronics, Kanagawa Institute of Technology
matumoto@ic.kanagawa-it.ac.jp

Abstract

This study develops an automatic method to estimate the degree of concentration on a task mainly focusing ones in a classroom. To know the degree is useful to understand a situation of a student and to improve the qualities of a classroom. Concentration depends on hidden internal states of a human so that it is difficult to measure the degree from observable external his/her behaviors. In this study, we first assume concentration degrees can be characterized by a set of activity patterns that is defined over each task. Thus we first identify activities, each of which completes in a short time interval, using data obtained from the standard accelerometer in a smartphone, and next the degree is calculated based on patterns of them. From another point of view, a historical record of calculated degrees is a part of important information for the lifelog of a person. We then periodically compose a tweet of him/her, and post it over Twitter. We show how we can use the lifelog in effective classroom education. By investigating tweets in a classroom, this system scores points for students, and it further arises competition sprits of students with a use of gamification. This poster presentation shows an outline of the system, primary implementation, and some experimental results.

Keywords: activity recognition, education system, data mining, lifelog, SNS, Twitter,

1 Introduction

Intelligent information systems are effective in classroom education, and in fact many lecture support systems are already developed and can be used in a classroom. In particular, several studies show the applicability of the SNS including Twitter approaches, whose main advantages are possibilities of connections among students and comments on mutual posts. We also attach importance to a smartphone as an effective ubiquitous educational tool, where most of university students now wear it almost all the time. We especially use a smartphone as a device for recognizing activities of a student. Our approach in this study combines these two recent technologies, recognizing a student activities based on sensing data with a smartphone and posting the activities over SNS as a part of his/her lifelog. In previous studies using SNS, a student has to post short messages by oneself. This cost cannot ignorable in general and be a tougher task for continuing over a long term. On the hand, our approach is based on an automatized posting so that the cost decreases to almost ignorable level.

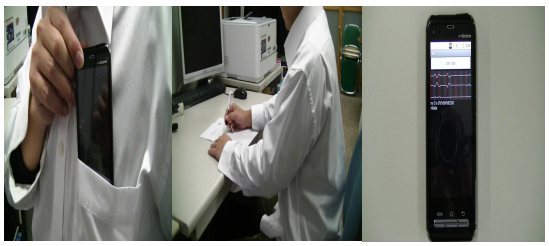


Figure 1. Using Smartphone

2 Activity Recognition using Smaptphone

Figure 1 shows a smartphone worn in a breast pocket of a student. We assume the student keep wearing it in the same pocket during a classroom, and the smartphone continuously collects sensing data of the accelerometer, the sensing frequency is about 100Hz a second. The similar approach is already used in [2].

Activity recognition is carried out by using the collected data and predefined activity template pattern data. This task recognizes an activity every 2 seconds, where the basic activities listed below are completed in this short interval. The activities recognized here are i) sitting lazy; ii) shaking body; iii) looking back; iv) unstably moving; v) keeping quiet.

The method used in the activity recognition is based on the idea of time-series data datamining [1] with clusterings and decision trees.

After extracting feature patterns, which are characteristic sub-sequences of data, and then build a decision tree to classify a given time-series data into the activities above. Feature patterns depend on persons so that each student has own feature patters and own decision tree. In other words, we need a learning task collecting sensing data and completing datamining for each student.

2.1 Calculating Concentration Degree

The key point of this system is a calculation of a concentration degree based on recognized activities. Concentration is a matter of unobservable brain activities. We then need to calculate its degree from body activities but there is no general method to do this. We develop a mechanism that is restrictedly applicable in the case of having a lecture sitting at a desk in a classroom. The basic idea here is based on observations that a student moves less and quiet when he/she deeply concentrate in a lecture. The activities listed above are recognized every 2 seconds, concentration degree is calculated every 10 seconds by considering a sequence having 5 activities. Experimental results including the precision of this method will be shown in the poster session.

3. Using SNS and Lifelog

Making a personal lifelog is a social activity to record various aspects of his/her life from viewpoints of events, situations, statuses, relationships, etc. The recorded lifelog is used to look back his/her life and can be used to discover unrealized aspects of the life, to reflect own activities, to decide actions based on precedents and to improve everything to reach better life. Moreover having a lifelog, which is open to a community, means to be a part of community and to extend strong connections among members. Members each other can comments on a record, a tweet in the case of Twitter, this naturally stimulate mutual competition in a classroom.

3.1 Automatic Tweeting

A good lifelog requires frequent records, posting many articles or tweets however costs time and patience. In particular it is difficult to post a tweet in a classroom with keeping enough concentration to a lecture. To solve this problem, we adopt an automatic posting. Calculating concentration degree is translated into a short message tweet and is posted over Twitter as we show in Figure 3.

Figure 3. Example of Tweets

3.2 Competing over Lifelogs

Lifelogs are shared over a classroom community. We then introduce an idea gamification [3] to increase self and mutual competitions to increase situations in classrooms. Awards or points obtained in the gamification increase motivations of students. The design and experimental results will be also shown in the poster session.

4. Conclusions

We propose an intelligent classroom information system that activates useful interactions and increases mutual understanding. The key idea is an automatic posting of concentration degree as a tweet, which is calculating based on sensing data with a smartphone. By investigating tweets in a classroom, this system scores points for students, and it further arises competition sprits of students with a use of gamification. Several experimental results with future problems will be shown in the poster session.

References

[1] Dong, G., and Pei, J, Sequence Data Mining. Springer, New York, 2007.

[2] Kwapisz,J.R.,Weiss,G.M.,and Moore, S.A., Activity Recognition using Cell Phone Accelerometers. *SIGKDD Explorations*, Vol.12, No.2, pp. 74-82, 2010.

[3] Muntean,C.I., Raising Engagement in E-learning through Gamification, ICVL, pp.323-329.

IT Governance in Shared Services in Public Sector

Omar M. Ali
School of IT-Faculty of Business and Law-University of Southern Queensland
Omar.Ali@usq.edu.au

Abstract

The importance of enterprise governance in IT has been demonstrated in both the private and the public sector. Some of the issues that implementing IT governance addresses are IT-business alignment, the demonstrable business value of IT, effective IT portfolio management, maximisation of the value of IT investments as well as mitigating the risk of IT failure. As the pressure on government ICT departments grows to provide additional services, to respond quickly to new business requirements, to increase effectiveness and to continue to demonstrate business value while under increased scrutiny and tight budgets, relatively new models of service delivery are being considered and implemented in the public sector across the world, these include outsourcing, public-private partnerships and the creation of shared service centres (SSC) to deliver inter-departmental services to government agencies. This paper will focus on SSCs in the public sector and will review the literature and present a case study of IT governance failure in an SSC in the Queensland government's implementation of a centralised HR and payroll system. Based on the analysis completed, we explore the risks and benefits in SSCs and investigate key governance-related concerns in the creation, transition to and operation of SSCs in the public sector, highlighting the importance of the relationship between the SSC governance model and the scope of service delivery.

Keywords: IT Governance, shared services centres and public sector.

1 Introduction

IT has become a commodity, as Carr states in his seminal article [16]. Just like other commodities, its status as a commodity has paradoxically led to its increased importance in our everyday lives. This has transformed technology in both private enterprise and government from a curiosity, a silver bullet to an essential tool upon which our everyday lives depend. After the dot com bust, the Y2K con and the global financial crises, IT is expected to constantly prove its business value. Spectacular failures such as the one explored in the case study should not happen, however both the private and public sectors are littered with examples of such costly failures.

IT governance, as opposed to IT management which deals with supply of IT services, is meant to control the demand of IT, by ensuring IT-business alignment, formulating the business requirements of IT and ensuring effective use of investments in technology.

With the advent of e-government initiatives, the public sector has also become increasingly dependent on IT. The famous Gershon report promoted major cost savings in the UK, and it prompted Australia to commission a similar report. The public sector around the world started imitating the private sector by creating shared service centres [16].

Shared service centres present additional IT governance issues. How to ensure strategic alignment between government departments which historically have been independent while at the same time ensuring cost savings through economy of scale? How to look at IT governance from a 'meta-organisational' perspective? While we accept Grembergen & Haes's model of governance, the reality of applying this to a multi-tenanted environment is much more difficult in practice [6].

We believe spectacular failures present an opportunity for both academics and practitioners to gain valuable insight, identifying causes and working to prevent them in the future.

In this paper we have selected the failed HR and payroll implementation in the Queensland Health Department. We will first analyse the failure itself and then conduct research in the literature to compare the results and extrapolate our conclusions on the ideal configuration of a shared service. Since, in our view, enterprise governance of IT must

formulate the demands of IT, when establishing these demands, it is important to clearly understand the requirements, to evaluate the inherent risks and to understand the limits of the feasible when embarking on the long-term strategic decision to create a shared service centre. This paper will attempt to identify success factors in SSCs in the public sector.

2 Literature review

The topic of shared service governance in the public sector is a complicated topic; as so few papers have dealt with this topic specifically. However, the papers we have selected are comprehensive and include relevant case studies.

The importance of clear decision rights and proposes governance structures considered optimal based on existing SSCs, normally three or four levels starting at national level with a governing board, then at executive level with a steering committee as well as an implementation office at this level because transformation projects in SSC environments can be particularly complex and require constant guidance. The third level, the managerial level, deals with partnership management among the members of the SSC as well as process and strategy review and SLA approvals. Sometimes this level is split even further into a fourth, tactical level which deals with day-to-day operations and user committees to gain user feedback on performance for a continuous service improvement cycle [5].

The paper also presents eight important processes to consider when designing governance for an SSC [5]. These are:-

1. Customer-focused strategic leadership.
2. Strategic alignment aided by separating strategic from operational concerns and avoiding charge-back models of cost recovery.
3. Risk management by clearly defining service delivery expectations across all SSC participants.
4. Performance management via metrics to prove the cost savings gained by SSC service delivery as well as issue resolution (the use of the Kaplan and Norton's balanced scorecard is recommended).
5. Control and legislative compliance, for issues such as privacy and accountability.
6. Relationship management, of vital importance in an SSC, which addresses transparency, communications to participants during the transition to the SSC and attentiveness to the cultural. impact upon participants that an SSC transition causes
7. Transformation management, including skilled staffing, defining transition timelines and managing the growth and scope of the SSC.
8. Value management, to deal with on-going funding for services (who pays for what and in what proportion?).

The significant viewpoints of the research governance on SSCs have two different viewpoints. One is the resource based view (RBV) and another dynamic capability view (DCV) [9]. RBV considers resources to be scarce and views an SSC as a way to achieve better efficiency in the use of resources by sharing services [9]. DCV instead concerns itself with the ability of an organisation to change in response to a rapidly changing environment and the changing needs of each stakeholder in the SSC [9]. Both the ability to achieve greater efficiency and respond to the needs of SSC participants (IT-business alignment) are essentially governance issues.

This paper stresses the importance of the three main aspects of governance: processes, structures and relational mechanisms [9]. The two case studies examined in the paper show two different governance structures for two different types of SSCs. In the first case, the SSC provides central, non-customized services to government departments in Holland. The driver for this centralised approach is to avoid duplication of efforts for commoditized ICT services. It has a centralised, top-down decision-making structure, with representatives from the participants being senior executives with more strategic than operational functions within their organisations. The services are offered and some later discarded based on the uptake. Account management focuses mainly on SLA negotiation and for escalation of performance issues. The central governance board focuses on risk and project management, funding model reviews and scope expansion.

The second case study is instead, a decentralised, federated model with decision-making based on participant consensus. In this case, the SSC was essentially a merger of the ICT departments of multiple municipalities in

Holland. The SSC was given the mandate to provide all ICT services to its participants. New requirements, performance issues, feedback and project and prioritization management are governed by user boards comprised of line managers, process owners and administrative users. The SSC governance board in this case has a compliance and as well as decision-making rights on major architectural choices [2].

The paper conclusions trend with the following key sentence 'The level of commoditisation of shared services determines the efficiency that can be gained; however, it limits the customisation and consumer orientation'. It concludes by underlining the importance of governance in SSCs and in balancing the governance mechanisms with the agreed objectives of the participants when creating the SSC. Governance will need to carefully balance customisation and commoditization to achieve efficiency while ensuring consumer orientation.

According to the important a study which conducted that the public sector organizations from several countries have embraced the idea of shared services [20]. The idea promises both high quality service and cost reduction [21] [4]. The idea is that of combining the best elements of decentralized and centralized service delivery. The adoption of common services in public sector organizations is based on a combination of process, people and policy changes [17]. They explore a successful HR and payroll implementation by four government agencies. In this case, senior management from each participant recognized the efficiencies to be gained and created an SSC with a federated governance model based on voluntary participation [21]. This SSC was so successful that a national board was created in Sweden to drive a whole-of government shared services initiative. However, this new national board encountered difficulties due to expert staff being unwilling to move to the new SSC and not seeing the larger picture of greater benefit for all participants. Also, a point of contention was due to a nation-wide SSC essentially reducing the administrative budgets to individual agencies, which caused some political resistance. The paper stresses the importance not just of senior management buy-in, but also buy-in at the operational level by individual employees. Employee resistance to an SSC transition is also exacerbated by the sense of life-long employment as a public servant, so that inertia/passive resistance by key experienced staff can be a severe impediment to organizational transformation initiatives [21].

Another interesting study that, conducted cost saving and efficiency as prime drivers for the creation of SSCs [22] [4]. SSCs facilitate cost reduction by introducing a transparent vendor for service delivery and as a concentration point of expertise. It implies a long-term strategy decision. However this paper highlights the complexity that such a move entails due to the migration of legacy systems as well as divergent expectations among participants which renders IT-business alignment in an SSC very difficult. He cites a case study of another failed HR/payroll system rollout in the Netherlands as an example and compares the bottom-up emergent process growth versus top-down steering committee direction setting governance structures. In the case study, the issues arose from a lack of a programme director and top-down decision making which ignored the pre-existing complexity of the HR system. The SSC in question took it upon itself to make unilateral requirement decisions and no clear migration strategy or roadmap was developed. This led to alienation and a lack of trust by the participants in the SSC. In turn, key integration projects with existing systems failed and the Minister of the Interior had to step in and install a programme management office. This failed to turn around the project and the SSC project was declared a defeat. Wagenaar [22] defines the scope of an SSC by:-

- Type and number of processes supported.
- The organisations to which the services will be provided.

An SSC can be compared to a particular form of outsourcing; however the paper identifies a number of governance risks in the creation of, transition to and operation of an SSC. These risks are mainly organisational and politically related, as existing management structures resist changes and power shifts to the SSC. The risks are listed as:-

- Resistance against one-size-fits-all.
- Resistance against power concentrations.
- Resistance against higher than expected initial costs associated with enterprise reorganization.

- Resistance to lengthy project time lines which may give the perception to users that no benefits are being achieved.

The paper also identifies the dilemmas of a public sector SSC, both in its design and in its implementation. For design aspects a prime concern is defining the scope of the SSC, as increasing its scope too much can have the effect of reducing overall efficiency gains. Choosing the geographical location of the centre can also be both a practical and a political issue. Relational mechanism design is also important, with transparency being a key factor. The relationships between a participant and the SSC, whether formal or informal, should always lead to transparent consensus-based decisions [7].

As for implementation dilemmas a prime one is decision-making, whether participants will delegate certain decisions to be made centrally, or whether a decentralised, federated governance structure is used. Expectation and communication management is also a dilemma.

The recommendation here is to run open and well-communicated pilot projects first, gain acceptance and then start a full project which need not be as open as the pilot.

The paper concludes with no 'easy answers', given that the topic is still too new to have developed governance design patterns and 'best practices'. The main issue in the public sector is the lack of a central board of directors, which makes decision-making difficult. The general recommendation given is to start with limited services, as the risks of implementing large, sophisticated services such as ERP and HR software are great and the performance gains are hard to evaluate.

Creating an SSC involves trade-offs and an effective governance structure [8]. An SSC is not easy to accomplish as it requires considerable changes in organisational arrangements, a coordination mechanism, business processes and responsibility allocation. It is also difficult to make optimal decisions since it is hard to quantify success factors. To be considered successful the SSC must achieve a delicate balance of service improvement, cost reduction and innovation [4]. Finally, outcomes are hard to measure as participants' perception [8].

Important aspects which cannot be overlooked include the design of the governance structure in relation to the existing participant structures. Architecture must be established that takes into account both the centralised and de-centralised elements. The right expectations must be set from the start and effectively communicated to stakeholders. Communication is key and must make clear to participants and end users the envisaged strategy as well as on-going progress and the benefits the SSC is achieving [8].

There is no evidence to support the optimal business model for an SSC that achieves all benefits in each distinct situation. The paper describes the following common types of SSCs [8]:

1. Central department for IT staff.
2. An internal joint venture.
3. Infrastructure facilities for multiple business units.
4. A service arm in a single department which provides services to others.
5. A separate service firm, similar to outsourcing arrangements with the participants as its customers.

According to important study which conducted that, shared service centres (SSCs) are becoming more popular as a service delivery model [11]. Currently, there are significant barriers to transforming structures and processes within local government. Some of the frequently mentioned problems include the inability to predict the outcomes of a change and the increased difficulty in achieving an IT-business alignment [18][19]. Shared services enable governments to focus resources on the primary high-impact activities at the core of their mission rather than on routine administrative functions. However, expectations of SSCs are often high, and sometimes not entirely realistic. Furthermore, it is difficult to realize all the objectives, because the introduction of an SSC often involves a series of complex, interrelated objectives and involves a large number of stakeholders. The paper highlights the need for decision support systems to simulate and evaluate the impact of shared service delivery.

3 IT governance case study

CorpTech is a shared service centre created by the Queensland government as part of a wider, shared services initiative which began in 2003. This initiative included whole-of-government delivery of payroll, rostering, purchasing, inventory management, asset management,

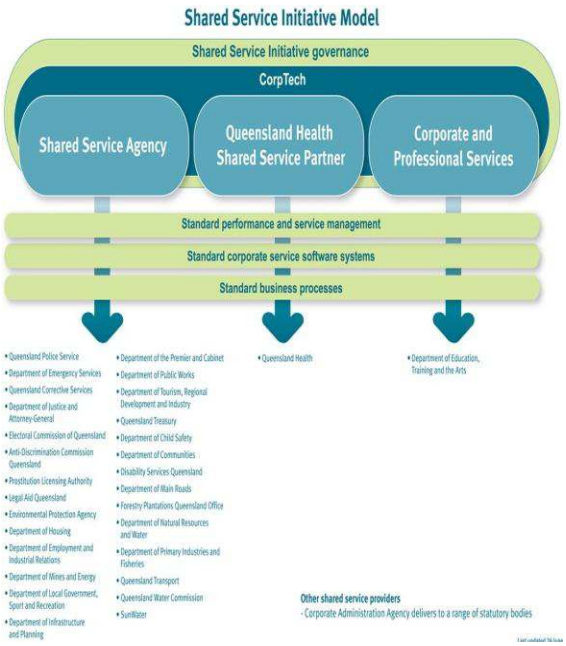
accounts payable and receivable [15]. CorpTech was made responsible for systems to administer finances and HR. CorpTech, in turn, established the Corporate Solutions Programme to implement a whole-of-government payroll, finance and HR system [15].

CorpTech initiated a project to roll out a centralised HR and payroll system to all major agencies across the public sector [13] [14]. The software selected by Corptech was SAP, which would also manage and direct implementation. Queensland health (QH) was originally scheduled to begin implementation of the system in 2006. Queensland health's legacy system LATTICE would no longer be supported by its vendor from July 2008. Delays in the whole-of-government initiative prompted QH to escalate and have its implementation schedule given precedence [13] [14].

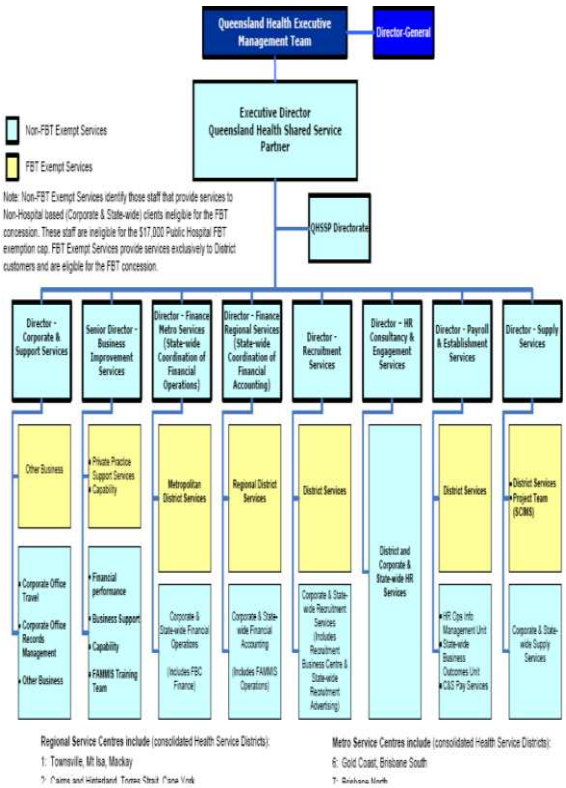
However there were unclear governance structures and decision-making rights between the main agencies involved, in this case Corptech, IBM the contracted vendor and the Queensland health department [13] [14].
The original project timeline set August 2008 as a go-live date, however, incorrect business requirements, a lack of a post-implementation test strategy, a lack of periodic reviews of the business requirements, lack of useability testing and validation of new business processes and a lack of a contingency plan cause the project to be delayed by almost two years [1].

After 47 business requirement changes and $64.5m later (300% over-budget), when the system did go live, a majority of the 76,000 employees at QH were paid incorrectly, superannuation contributions were incorrect and for the close of the financial year employees were instructed to calculate their own group certificates. The issue, at the time of writing, still has not been rectified and there have been calls for the QLD health minister to be sacked. Blame was also directed at the department of public works which controlled CorpTech. The QLD government also entered into litigation with IBM over the matter [1].

The spectacular failure of this project caused the Queensland government to review and effectively abandon its centralised, whole-of-government shared services approach to payroll and to order a review of its entire shared services initiative [1].



Shared service initiative model



Structure of QLD health department

4 Results

From the literature review, a number of enabling factors and barriers can be identified which are crucial to the successful design and implementation of an SSC. These are as follows [3]:

- Management commitment: This is crucial because without management commitment

to the SSC, it is doomed to failure from the outset.

- Communication and consultation**:** All stakeholders must be consulted so as to best determine the needs of various departments and accordingly create an implementation plan for the SSC. Constant communication at each step of the process is crucial.
- Centralized versus de-centralized system**:** The decision between a centralised system, a de-centralised system or a combination of both methods must be made prior to rolling out the SSC services. This will enable a suitable plan of action which is focused towards meeting the objectives of the organisation as a whole.
- Process, people and policy**:** These must be kept in mind at all times during the pre-planning and planning phases. The process must be explicitly defined for all to understand; the people's attitudes towards change must be gauged and accordingly these people must be informed thoroughly about the benefits of the new process/system so that they may be able to adopt them seamlessly and finally policies must be in place to facilitate smooth transition to the new shared services system.
- Appropriate services**:** Choosing the appropriate technology can make or break the shared services program. Such technology includes software and hardware.
- Change management**:** This is needed so that the migration or transition is done in an organised manner following a pre-defined setup plan.
- Administrative organisation**:** A clear hierarchical power structure must be defined for the new system so as to designate responsibility for various aspects of the shared services to specific individuals who will be held accountable.
- Stepwise implementation**:** Start small and test run the system to identify problems to be fixed before rolling out the program. Expand gradually using a realistic time frame. This way problems arising can be dealt with before moving on.
- Contingency/backup plan**:** This is required to forestall any major problems during the implementation of the shared services initiative.
- Monitoring and evaluation**:** This is required to determine if the goals of the organisation and the shared services initiative are being met or not. It serves as checks and balances system to ensure that the objectives of the program are being met.

5 Research Analysis

Based on the literature review and the information from the case study available to date, it is clear that shared services initiatives, while theoretically providing efficiencies and cost savings, are fraught with additional risks and complexities which IT governance structures, processes and relational mechanisms in public sector departments must address, not only before the start of such an undertaking but also by carefully measuring performance during the implementation and operation phases. From the case study, we identified the following governance issues:

- The SSC governance structure was not well defined, and, in fact, was re-shuffled mid-way through the project. The QLD government failed to take into consideration the additional governance complexity of a shared service model.
- While the department of health did make use of the balanced score card approach for measuring the value of IT investments, it is clear that this model was not modified to take into account inter-departmental projects such as the one studied.
- The government failed to prepare a contingency plan, and opted for a 'big-bang' approach to implementing a complex, highly-customized service such as payroll, and failed to appropriately test the system prior to the go-live date. Indeed, test criteria were lowered to allow the system to pass QA. This shows top-down political pressure. Thus it failed in appropriate risk management.
- The fact that the scope and requirements signed off by the government to the main contractor, IBM, were later found not to meet the business needs and the fact that both parties blamed the contractor shows the failure to set up appropriate programme management and processes for requirement analysis and sign-off.
- The massive cost blow-out shows that there was a lack of portfolio management to measure the business benefits of the large

investment and they failed to take appropriate steps once it was clear that the cost of the project had been massively under-estimated and failure was imminent.

The fact that the Queensland government, in light of this failure, ordered a review into the whole-of-government initiative shows they are having difficulty clearly demonstrating to stakeholders the cost savings and benefits of shared services. This is an issue highlighted in the literature review.

6 Recommendations

Based on the issues exposed in the case study and the literature review conducted, several recommendations can be provided.

- First of all, in the case of government departments, which historically have operated as independent units, the barriers to shared service adoption can be enormous and must not be underestimated. The literature show no unified best practice governance model for shared services in the public sector. Also, there will always be a trade-off between efficiency and service levels, and proposes the use of complex simulation models to evaluate the impact of shared services [11]. Stresses the importance of clear decision rights among the SSC and its participants as key, and also stresses the importance of stakeholder communication and transformation management, relationship management and performance management, including the use of Norton and Kaplan's balanced score-card at the SSC level, not just within individual participants' organisations [5]. The SSC must continuously prove its value and cost savings. This did not happen in the case study and led to the QLD government ordering a review of its shared services strategy.
- The researchers describe a centralised versus a decentralised SSC and state that the cost-savings of providing a shared service are directly related to the level of customisation required by each participant [10] [12]. In the case study example, proper cost-benefit analysis might have shown that the cost savings of a highly customised service such as payroll would not have justified the increased complexity and additional risks created by providing it as a shared service [10].
- Another researcher specifically singles out ERP and payroll applications as lacking the modularity necessary for providing them as a shared service. This paper provides a similar failed HR project where decision-making lines between the vendor, the participant and the SSC were not clear, leading to political issues and participants' alienation from the SSC. What is obvious both here and in the Australian case study is that the existence of an SSC in the project added an additional layer of complexity which was not addressed by clear governance structures, processes and relational mechanisms [22].
- Finally, we recommend that for the QLD case study, the SSC could have been used to aggregate demand and drive down the cost of payroll licensing software by providing a state-wide licensing software agreement. The SSC could also have provided central infrastructure and network services to host the payroll application under an infrastructure-as-a-service (IaaS) model. Both these actions would have immediately demonstrated the cost-saving value of CorpTech, leaving the actual complexities of implementation to each participant.

7 Conclusion

Creating an SSC involves trade-offs and an effective organisational and management structure is crucial. An SSC is not easy to accomplish as it requires considerable changes in organisational arrangements, coordination mechanisms, business processes and responsibility allocation. Obtaining the full benefits is very difficult and requires that the SSC capture the best elements of centralised and decentralised organisation. It is also difficult to make optimal decisions since it is harder to quantify success factors. To be considered successful, the SSC must achieve a delicate balance of service improvement, cost reduction and continuous innovation. Finally, outcomes are hard to measure as participants' perception. Important aspects which cannot be overlooked include the design of the management structure in relation to the participants' existing structures. The right expectations must be set from the start and effectively communicated to stakeholders. The scope of services to be delivered via the shared

services model must be carefully evaluated; the scope must then determine which governance model is appropriate. Infrastructural IT services are better suited to a shared service delivery model. Communication is the key and must make clear to participants and end users the envisaged strategy as well as on-going progress and the benefits the SSC is achieving.

We conducted this research in a limited timeframe. It is clear, however, that the topic of shared service centres in the public sector is still new and should be approached which caution. There is evidence from the other case studies which show that not all ICT services are fit for central delivery. Also, found that while an SSC can successfully provide some services at a reduced cost, it may not be appropriate to provide all types of services to its participants. Indeed, the key concepts found in the literature have been to start slowly and incrementally, constantly demonstrating the benefits at every stage. The Queensland approach here appears to have been top-down driven and the SSC scope was over-reaching. When the first signs of project delay and potential failure became, apparent bad decisions were made, most certainly due to political pressure and unclear governance structures. This clearly defies most recommendations in the literature. Future research points to comparing shared services with outsourcing arrangements and applying frameworks developed for evaluating outsourcing services to shared services.

References

[1] Anatomy of an IT Disaster - How the IBM/SAP/Workbrain Queensland Health Payroll System Project Failed 2010, accessed 22/09/2012, available at http://ddkonline.blogspot.com/2010/07/anatomy-of-it-disaster-how.html

[2] Becker, J, Niehaves, B & Krause, A 2009,'Shared services strategies and their determinants: A multiple case study analysis in the public sector', *Proceedings of the Fifteenth Americas Conference on Information Systems*, San Francisco, California.

[3] Borman, M 2010,' Characteristics of a successful shared services centre in the Australian public sector', *Transforming Government: People, Process and Policy*, vol.4, no.3, p.210.

[4] Dollery, B, Akimov, A & Byrnes, J 2009,'Shared services in Australian local government:rationale, alternative models and empirical evidence', *The Australian Journal of Public Administration*, vol. 68, no. 2, pp. 208-219.

[5] Grant, G, McKnight, S, Uruthirapathy, A and Brown, A 2007,'Designing governance for shared services organizations in the public service, *Government Information Quarterly*, vol.24, no.1, pp.522-538.

[6] Grembergen, W V and De Haes 2009, 'Enterprise governance of information technology', *Concepts of Enterprise Governance of IT*, Belgium, pp.1-20.

[7] Helbig, N, Hrdinová, J & Canestraro, D 2009,'Enterprise IT governance at the state level: An emerging picture', *The Proceedings of the 10th International Digital Government Research Conference*, Albany, New York.

[8] Janssen, M and Joha, A 2006,'Motives for establishing shared service centers in public administration', *International Journal of Information Management*, vol.26, no.1, pp.102-115.

[9] Janssen, M and Joha, A 2007,'Understanding IT governance for the operation of shared services in public service networks, '*International Journal of Networking and Virtual Organizations*, vol. 4, no. 1, pp. 20-34.

[10] Janssen, M and Joha, A 2009, 'Public-private partnerships, outsourcing or shared services centres?', *Transforming Government: People, Process and Policy*, vol. 4, no. 3, pp. 232-248.

[11] Janssen, M, Joha, A and Zuurmond, A 2009,'Simulation and animation for adopting shared services: Evaluating and comparing alternative arrangements', *Government Information Quarterly*, vol.26, no.1, pp. 15-24.

[12] Jereb, C, Kuchem, R and Sohn, W 2009, 'Shed light on the black box: how Deutsche Post DHL achieved global transparency over payroll processes', *Strategic HR Review*, vol. 8, no. 5, pp. 23-31.

[13] Payroll saga: Bligh admits 'fundamental failures' 2010, accessed 02/10/2012, Brisbane times.com.au.

[14] Qld health IT disaster shakes up Australia's government CIOs 2010, *CIO*, accessed 05/09/2012, cio.com.au/…/qld_health_it_disaster_sha

[15] Qld government 2007, accessed on 17/02/2012, available at http://www.treasury.qld.gov.au/knowledge/docs/annual-reports/200708/outputs/corptech.shtml

[16] Schulz, V and Brenner, W 2009,'Characteristics of shared service centres', *Transforming Government: People, Process and Policy*, vol. 4, no. 3, pp. 210-219.

[17] Steven De Haes, S and Grembergen, WV 2005,' IT governance structures, processes and relational mechanisms: Achieving IT/business alignment in a major Belgian financial group', *Proceedings of the 38th Hawaii International Conference on System Sciences*' pp. 1-10.

[18] Steven De Haes, S and Grembergen, WV 2008,' Analysing the Relationship between IT Governance and Business/IT Alignment Maturity', *Proceedings of the 41st Hawaii International Conference on System Sciences*' pp. 1-10.

[19] Turle, M 2010,'Shared services: An online of key contractual issues', *Computer Law and Security Review*, vol.26, no.1, pp. 178-184.

[20] Ulbrich, F 2009,'Adopting shared services in a public- sector organization', *Transforming Government: People, Process and Policy*, vol. 4, no. 3, pp. 249-255.

[21] Ulbrich, F 2010, 'Adopting shared services in a public-sector organization', *Transforming government*, vol.4, no.3, pp. 249-265.

[22] Wagenaar, R W 2006,'Governance of shared service centers in public administration: dilemma's and trade-offs', Delft University of Technology, The Netherlands.

e-Government and its implementation Challenges in Nepal

Subarna Shakya
Department of Electronics and Computer Engineering
Central Campus, Institute Of Engineering
Tribhuvan University, Nepal
e-mail:drss@ioe.edu.np

Abstract

e-Government is the use of Information and Communication Technologies to promote more efficient and effective government, and make it more accessible and accountable to the citizens. The e-readiness index of Nepal is found to be low as compared to other countries. There are various challenges for the implementation of e-government in Nepal. These challenges are like low literacy, low per capita income and limited financial resource. In this paper a conceptual framework is suggested for the effective implementation of e-government in Nepal. The conceptual framework can be further validated in the real life situation.

Keywords

e-Government, e-readiness, low literacy, Low per Capita, conceptual framework

1. Introduction

e-Government is the use of information technology to provide citizen and organizations with more convenient access to government information and services and to provide delivery of public services to citizen, business partners, and those working in the public sector. In other words e-Government is the use of information and communication technologies (ICTs) to improve the activities of public sector organizations. e-Government has very significant role in the development of a nation. It has brought revolution in the governance of the government. According to Economist Intelligence Unit the e-readiness index of Nepal is low. The implementation of e-government is very challenging because of the low level of literacy, low per capita income and insufficient infrastructure for the implementation of e-government in Nepal.

2. E-government Project in Nepal

The government of Nepal with support of the Korea IT Industry Promotion Agency (KIPA) prepared an e-Government Master plan in November 2006[1]. In order to establish the foundation for investment phase of the Master Plan, The Asian Development Bank (ADB) provided a project preparatory technical assistance (PPTA) to the Government of Nepal [2]. According to the PPTA, the implementation plan was initiated and ADB agreed to invest grant fund for Nepal and after that the government of Nepal and Asian Development Bank (ADB) signed the grant agreement of ICT development project dated 23 May 2008[3]. Then after e-Government implementation project starts in Nepal. The Project involves[5] (i) modernizing rural communities, particularly in remote areas, by improving rural connectivity through wireless broadband networks, mobilizing community socioeconomic activities through village network portals, and building telecenters to improve last-mile access to services in remote rural areas; (ii) building a government information and communication technology (ICT) network, which allows government-to-government exchange of data and information and central management of government data and information; (iii) developing various priority e-Government applications; and (iv) developing and implementing human resources development programs.

There have been e-Government initiatives in the country at national level. The program will lead to socioeconomic improvements in remote and rural communities through ICT and improved ICT uptake in the community, business, and government. To this end, the program will (i) make ICT more accessible, affordable, inclusive, sustainable, and useful to remote and rural communities; (ii) make public services more citizen-centered and business-friendly through ICT; (iii) improve accessibility, efficiency, and transparency in government service

delivery through ICT; and (iv) enhance ICT business and industry. The priority project is as follows [4]:

Rural e-Community

The Project will modernize rural communities, particularly in remote areas, by improving rural connectivity through wireless broadband networks in districts, by mobilizing community socioeconomic activities through a village network portal that will allow villagers to share their social capital, and by building telecenters to improve last-mile access to services in remote rural areas.

Government Network

The Project will build a government ICT network, which allows government-to-government exchange of data and information, and central management of government data and information, with suitable protection and efficient backup and recovery provisions. The government ICT network component will comprise (i) the establishment of government information and data center, and (ii) the development of government groupware.

e-Government Applications

36. The Project will develop various priority e-government applications that (i) are priority applications; (ii) are in areas requiring few legislative changes; (iii) provide easily identifiable benefits; (iv) respond to owners' keenness to initiate change and own the application; (v) are relevant to the immediate needs of the community; (vi) hasten the achievement of MDGs; (vii) answer actual or latent demand, cost, and organizational capability; or (viii) are likely to have quick success, with significant dividends to the community.

Human Resources Development for e-Governance

The lack of adequate knowledge and skills in ICT and e-Governance in the public sector has given rise to (i) poor-quality or nonexistent services; (ii) persistent corruption due to lack of transparency; (iii) unnecessary duplication and inefficiency in government work; (iv) flawed decisions because of a lack of information and analysis; (v) coordination difficulties and resistance to innovation; and (vi) poor use of financial, human, physical, and technical resources. To address this capacity problem, the Project will (i) build awareness, knowledge, and skills in ICT governance among key stakeholders to help improve efficiency in the delivery of e-services to the community; (ii) establish computer laboratories for the capacity development of institutions promoting ICT human resources development, and strengthen networking between training institutions and support for functional linkages; (iii) revise the training curriculum and develop new curricula for public training institutions, to improve the quality of the training curriculum and materials in ICT governance and enhance the cost-effectiveness of training programs by developing high-quality training materials; (iv) share knowledge and experiences to promote ICT governance and applications through exchange and fellowship programs designed to recognize the contributions and commitment of exemplary civil servants and NGO leaders to the effective and efficient adoption of ICT governance applications; and (v) support the development of ICT governance courses, a new curriculum for ICT governance, and new teaching materials in ICT governance, in association with universities and research institutes.

3. Implementation Arrangements

The Office of the Prime Minister and Council of Ministers (OPMCM) [4] will be responsible for overall project implementation and coordination as the Executing Agency. A project management unit (PMU) will be established in the OPMCM, within one month of effectiveness. It will be headed by a chief project director (secretary), who will be supported by two project directors, one for administration (joint secretary) and another for technical aspects (NITC executive director), and a treasurer, as well as project management consultants, who will provide technical inputs for the appropriate supervision of project implementation. The chief project director will have overall responsibility for project management. Each component and subcomponent of the Project will be assigned an implementing agency as follows: (i) for the rural e-community component, the Ministry of Information and Communications (MOIC); (ii) for the government network component, the Ministry of Environment, Science and Technology (MOEST); (iii) for the enterprise architecture subcomponent, HLCIT; (iv) for the national identification system subcomponent, the Ministry of Home Affairs; (v) for the e-governance in the Public Service Commission subcomponent, the Public Service Commission; (vi) for the land records management subcomponent, the Ministry of Land Reform and Management; (vii) for the vehicle registration and driver's licenses subcomponent, the Ministry of Labor and Transport Management; and (viii) for the human resources development component, the Ministry of General Administration.

Project implementation units (PIUs) will be named by the various Implementing Agencies for their respective components and subcomponents, as follows: NITC, for

the government network component and the government representation portal subcomponent; the Department of Land Reforms and Management, for the land records management subcomponent; and the Department of Transport Management, for the vehicle registration and driver's licenses subcomponent. For other components and subcomponents, the Implementing Agency will establish a PIU headed by a project manager. The PIUs will be established within one month of effectiveness and will (i) coordinate with the contractors and monitor the day-to-day progress of project management and implementation, (ii) prepare fund withdrawal applications and submit them to the PMU under the OPMCM for endorsement, (iii) prepare monthly project progress reports and submit them to the PMU under OPMCM, (iv) take charge of procurement for their respective areas of assignment, and (v) keep the project accounts and prepare the annual reports.

Overall project implementation will be supervised by a project steering committee (PSC). The PSC will be established within one month of effectiveness, will have the chief secretary of OPMCM as chair, and will comprise the project director (vice-chair), the secretaries of MOIC and MOEST, the member-secretary of HLCIT, and the executive director of NITC (committee secretary). The PSC will meet regularly at least twice a year, or more often if necessary, to set the direction for overall project implementation, guide the PMU on specific project implementation issues, and resolve any dispute between the PMU and the project implementation agencies.

4. Nepal's Position on e-readiness

e-Readiness is the ability to use Information and Communication Technologies (ICT) to develop one's economy and to foster one's welfare. Each year, the Economist Intelligence Unit produces a ranking of e-readiness across countries, based on six pillars of e-readiness: connectivity & technology infrastructure, business environment, social & cultural environment, legal environment, government policy & vision and consumer & business adoption. Nepal is at 153 positions with e-readiness in 2012[5].

Table 1: e-Government readiness rankings for Southern Asia

(Source: United Nations e-Gov. Survey 2012, Page 26, Table 117)

5. Challenges for implementation of e-government in Nepal

	e-Gov. Development Index		**World e-Gov. development ranking**	
Country	**2012**	**2010**	**2012**	**2010**
Maldives	0.4994	0.4392	95	92
Iran	0.4876	0.4234	100	102
Sri Lanka	0.4357	0.0.3995	115	111
India	0.3829	0.3567	125	119
Bangladesh	0.2991	0.3028	150	134
Bhutan	0.2942	0.2598	152	152
Pakistan	0.2823	0.2755	156	146
Nepal	0.2664	0.2568	164	153
Afghanistan	0.1701	0.2098	184	168
Sub Region Average	0.3464	0.3248		
World Average	0.4882	0.4406		

Implementation of e- Government has changed the way of living of the people in many countries. However, in Nepal the implementation of e-Government is little difficult because of its developing status. The government agencies find lot of difficulties in the smooth implementation of e- government in Nepal because of low literacy, low per capita income, insufficient infrastructure and limited financial resource.

5.1 Low literacy

Literacy is defined as the ability to read and write with understanding in any language. A person who can merely read but cannot write is not classified as literate. Any formal education or minimum educational standard is not necessary to be considered literate. Georgia is the country having 100% literacy rate and at 1st rank in the list of literacy (Table: 2). Literacy level of Nepal is 48.6% and ranked 162 in the list and Mali is having literacy level 24% , is the lowest level in literacy level (176).

Table 2: Literacy rate of select countries.

S.N.	Country	Literacy Rate	Rank
1	Georgia	100.0	1
2	Cuba	99.8	2
3	Russia	99.4	12
4	Australia	99.0	18
5	Canada	99.0	18
6	Germany	99.0	18
7	Japan	99.0	18
8	Switzerland	99.0	18
9	United Kingdom	99.0	18
10	United States	99.0	18
11	China	99.9	86
12	Sri Lanka	90.7	87
13	Swaziland	79.6	124
14	India	61.0	147
15	Pakistan	49.9	160
16	Nepal	48.6	162
17	Mali	24.0	176

(Source: http://en.wikipedia.org/wiki/List_of_countries_by_literacy_rate)

5.2 Low per capita income

Per capita income means how much each individual receives, in monetary terms, of the yearly income generated in the country. This is what each citizen is to receive if the yearly national income is divided equally among everyone. Per capita income is usually reported in units of currency per year. Globally Bermuda has the highest per capita income followed by Luxembourg. India is at 160th rank in with 530 US$ average per capita income (Table 3).

Table 3: National Average Per Capita Income of select countries

S.N	Country	Per Capita Income in USD	Ranking
1	Bermuda	N/A	1
2	Luxembourg	43,940	2
3	Switzerland	39,880	4
4	United States	37,610	5
5	Japan	34,510	7
6	United Kingdom	28,350	12
7	Germany	25,250	22
8	Canada	23,930	24
9	Australia	21,650	27
10	Swaziland	1,350	127
11	China	1,100	133
12	Sri Lanka	930	140
13	India	530	160
14	Pakistan	470	166
15	Bangladesh	400	174
16	Nepal	240	192
17	Ethiopia	90	208

(Source: http://www.success-and-culture.net/articles/percapitaincome.shtml)

5.3 Limited financial Resource

The Gross Domestic Product (GDP) is one of the measures of national income and output for a given country's economy. GDP is defined as the total market value of all final goods and services produced within the country in a given period of time (usually a calendar year). GDP of a country is the measure of its financial strength. United States is having the highest GDP in the world followed by Japan, Germany, China, United Kingdom (Table 4) and India with GDP 1,237,000 million US $ is at rank 12th (Table 4). Nepal with GDP 12,640 million US $ is at rank 121st (Table 4).

Table 4: GDP of select countries in 2008 given by the CIA World Fact book

S.N	Country	GDP(Million USD)	Rank
1	United States	14,330,000	1
2	Japan	4,844,000	2
3	China	4,222,000	3
4	Germany	3,818,000	4
5	France	2,978,000	5
6	United Kingdom	2,787,000	6
7	Russia	1,757,000	8
8	Brazil	1,665,000	10
9	Canada	1,564,000	11
10	India	1,237,000	12
11	Australia	1,069,000	14
12	Switzerland	492,600	22
13	Pakistan	160,900	49
14	Sri Lanka	42,160	78
15	Nepal	12,640	121
16	Kiribati	71	189

(Source: http://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal))

6. A Strategic framework for implementation of e-government

According to study of e-readiness in Nepal and found the implementation challenges of e-Government in Nepal. A conceptual framework is suggested for the effective implementation of e-government in Nepal. The five stages of the conceptual framework is suggested as follows:

Vision and Action plan for e-government implementation

In the first stage the vision and action plan for the effective implementation of e-Government has to be determined. In this level it is planned that to what extend the e-government can be implemented.

Framework for e-readiness Assessment

In this second stage to fulfill the vision the e- readiness of Nepal should be assessed. In this stage compared with respect to other countries. The e- readiness reveals the position of Nepal with respect to the other countries.

Overcoming challenges of e-Government

In this stage, the assessments process the challenges for effective implementation of e-Government will be exposed.

These challenges are low literacy levels, low per capita income and limited financial resource in Nepal. The challenges should be overcome for the effective implementation of e-government.

Developing the environment for e-Government

In this stage to developed positive environment needs to meet the vision of e- government implementation. This environment is internal and external environment.

Implementation of e-Government

After that the e-government should be implemented. This is the final step of the conceptual framework for e-Government implementation.

A conceptual framework, which is shown in figure 1. This is a conceptual framework and can be validated in the real life situation.

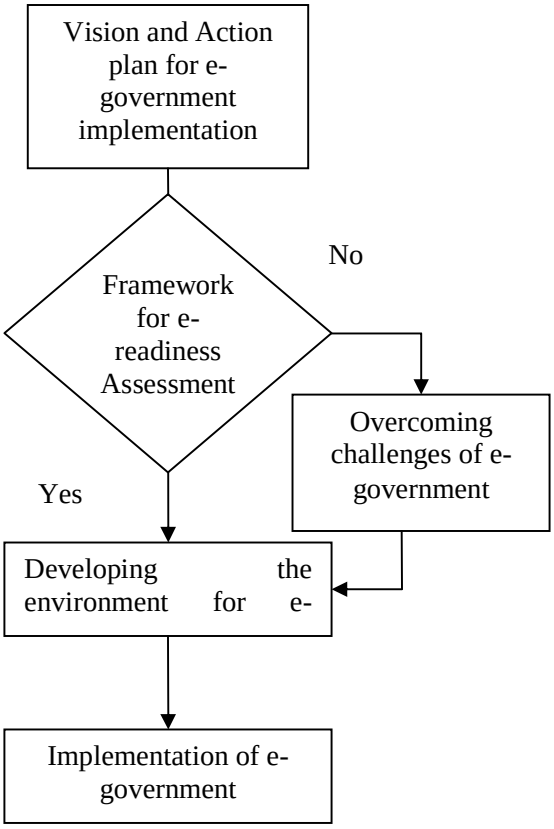


Figure 1: A conceptual framework for e-government implementation in Nepal

7. Conclusion

Thus Economist Intelligence Unit depicts that the e-readiness index of Nepal is low. According to this study the implementation of e-Government in Nepal found various challenges. The main challenges are like low literacy, low per capita income, and limited financial resource. A vision and leadership is required to implement the e-Government in Nepal. To meet the vision the challenges in the implementation of e-government should be overcome and needs to be developed environment for the effective implementation of e-government. A conceptual framework is developed for the effective implementation of e-Government in Nepal This conceptual framework and can be further validated in the real life situation.

8. References

[1]e-Government Master plan consulting report for Government of Nepal

[2] TA No. 4833-NEP: Preparing the Information and Communications Technology Development Project, a Programmme of Modernisation of Nepal, July 2007.

[3]http://www.adb.org/Documents/Legal-Agreements/NEP/38347/38347-NEP-GRJ.pdf

[4]ADB , Report and recommendation of the President to the Board of Directors, proposed Asian

[5]United nations e-Government Survey 2012

[6]http://en.wikipedia.org/wiki/E-readiness accessed during September 2008

[7]http://en.wikipedia.org/wiki/List_of_countries_by_literacy_rate) accessed during September 2008 Development Fund Grant Nepal: ICT development Project, 2007, pp.,7-11, 13

ICT Adoption by Indian Women: An analysis of two cases

Anindita Paul† **Deepa Ray‡**

†Indian Institute of Management Kozhikode
‡SP Jain Institute of Management and Research Mumbai
aninditapaul@gmail.com, deeparay@gmail.com

Abstract

This paper provides an analysis of two individual cases of women. The context is set in the southern Indian state of Kerala. The rich data gathered from at least an hour of interview and follow up enquiry are examined to understand the factors of adoption and use of ICT as well as how ICT has impacted their information behavior. The analysis was done using the UTAUT framework of technology acceptance and using a grounded information seeking approach that looks at emerging themes of ICT impact in the women's daily lives. Findings indicate changing perceptions of women's ICT orientation and use. Factors such as social influence, performance expectancy and effort expectancy play a role in ICT adoption and use. Also ICT has enabled the women to confidently take charge of their information needs in various circumstances.

Keywords: ICT, Women, Information Behavior, Information Systems, Technology Adoption

1 Introduction

Women's information seeking behavior and the support provided by Information and Communication Technologies (ICT) is an area that needs attention in today's dynamic ICT environment. ICT refers to any product that will *store, retrieve, manipulate, transmit or receive information electronically in a digital form*[1]. Women's use of ICT is of interest to information scientists who would like to look at aspects of information behavior (IB) as well as technology adoption. IB researchers have identified the need for more women-centric studies with ICT and have suggested directions for it [1, 2]. ICT has become a central part of today's lifestyle considering that most of the information accessed in the present times relies on various forms of ICTs such as computers, mobile phones, television etc. Women in the Indian society traditionally had restricted access to information owing to the lack of exposure to the external world. Additionally it's a common notion that women are not comfortable with technology use. As a result, understanding how women have adapted to ICTs as well as how ICTs have changed her information behavior is important.

From the Information Systems (IS) perspective, it is important to understand the role of ICT support for the tasks and activities that women need to perform. This will play an important role in technology adoption and help understand why women have been generally considered inept in their use of technology [1, 3].

Women today play multiple roles where they make various decisions about family health, children's education, running the household in a limited budget and so on. They need to manage their personal, household matters as well as professional responsibilities. In such scenarios, ICTs can play a vital role in helping women balance and fulfill their responsibilities. Women's information seeking behavior is driven by the need to take charge of her personal as well as professional goals and includes how she adopts/uses ICTs. ICT enable easy access to wide range of information. Not many studies have tried to understand women's adoption and use of ICT especially in the Indian context. In light of the increasing use of ICT in India it is important to understand how ICT is affecting the lives of groups or community of users.

This paper examines in-depth two cases of ICT use by two Indian women. Owing to the page limitation the benefits of case study research methodology can be found elsewhere in the literature. The two questions the case analysis is examining are- 1) the factors for ICT adoption and use and 2) how ICT impacts women's information behavior. The study can provide inputs to user-friendly services through ICT and wider adoption of ICT that can increase the ROI

[1] http://tutor2u.net/business/ict/intro_what_is_ict.htm

for businesses in the long run.

2 Literature Review

The IS area is rich with research on individual acceptance and use of technology. The UTAUT model [4] posits that Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI) and Facilitating Conditions (FC) are the factors that impact any individual's intention to use a particular technology and their actual use. Studies using the model have shown that gender differences play an important role in technology acceptance and use [5, 6]. Research on gender diversity and how it impacts ICT use has also received some attention in the literature [7].

For example, in one of the studies [5] conducted using the UTAUT model significant differences were found between men and women in their decision process to accept and use technology. Researchers found that in men, the attitude towards technology was more important and it determines their intention and actual use. However, women were significantly more influenced by their perception of how easy or difficult the use of technology was. Another important factor that influenced women's acceptance and use of technology were subjective norms, which included social and peer pressure.

However, most studies using the UTAUT model have been done in the western context. For example, Venkatesh et al. [5] called for research to study the UTAUT model in the context of developing countries such as India. Also, most technology adoption and use studies have looked at professional use, limited to workplace environments. In a 2012 study [8], the UTAUT model was extended to understand acceptance and use of technology from the consumer aspect .

IB of individuals in different cultures is constantly being shaped by emerging solutions. Women and ICT studies have addressed topics of women empowerment and feminism [3, 9], women and contribution to economy [10, 11], ICT use in women's health information seeking [12]. Ghobadi and Smith [9] cited a number of factors that influenced women's use of ICT such as fear of technology, support for safe cyber use as well as strategic use of technology. Some studies also looked at environmental influences when studying women's association with IT in formal and informal settings [13-15]. In a recent special issue of the 'Interaction with Computers' journal, the topic of feminism in Human-Computer Interaction was addressed [16].

Studies have also looked into the relationship between women and IT as a profession and posited that negative stereotypes led women to strive more to prove their worth and environmental context played an important role [17]. Further studies in IB [12] looked at aspects of women's health information seeking and found that they had a low awareness of resources that provide health and medical information and they apply weak search strategies to retrieve health information using ICT.

Women are considered to be more cautious and conservative about choosing search sources opting mostly for known sources or ones that are considered credible compared to men who are extremely self-confident and determined. Also, women access as many resources as possible including primary, secondary and tertiary as well as internal and external to their institute (formal and informal sources) [18].

3 Method

Two cases of urban women were explored, for purposes of better clarity [19]. The two urban women selected were professionals between 32-35yrs. Extensive data were collected through multiple-interviews with each of the two. Subjects for both the cases are of Indian nationality who resides in the city of Calicut in the state of Kerala in Southern India.

Data was collected from Oct 2012 to Jan 2013. The primary mode of the collection was through interviews. The interviews with the two subjects were open-ended with the first interview being face-to-face and rest as follow-up focused interviews. An interview guide was used for the interviews. Besides the data from the interviews, the subjects were asked to maintain an account of the typical things that they use ICT for in their daily lives. The researcher called them later to get the usage reports.

4 Findings and Discussion

Cases C1 and C2 were analyzed in the Indian social and environmental context. The analysis revealed aspects of ICT use in the women's everyday lives. As was apparent from the interviews, the two women considered ICT as an in-

separable part of their daily lives. However, there were differences in ICT use between the two cases.

C1 was a heavy user of iPad and C2 mostly used her laptop for information seeking and her smart phone to simply make calls. It was interesting to see the differences in the utilization patterns of C1 and C2, in spite of similar traits such as inclination to ICT. Both women had been using information technologies since the early days of its introduction in their respective townships and there was considerable support and encouragement from their families as well. C2 stated that she used to be the "local repairman" of her house during her childhood. SI (social influence) seemed to play a major role for both cases. Use of ICT for C1 started off with emails and that of C2 for preparing for entrance exams.

In case of C1, it was evident from the number of ICT devices she owned, its use and also her ICT related views, that she was pro-ICT (she owned a PC, iPad, iPod Touch and a Samsung Galaxy Y Duos S6102 smart phone). She thought IT had tremendous market offering and that it could not be ignored – *"It makes so much more sense to invest money and time on IT."* From the technology acceptance and use point of view, it could be said that her PE from ICT was very high. She relied on the Internet and applications on her devices for various needs from organizing, note taking to gym training etc.

C1 used her iPad the most. It was interesting to note that C1 did not use the laptop, which was C2's most commonly used device. On the other hand, C2 did not use her mobile device as much as C1 did. It could be inferred from the ICT use habits of C1 and C2 that both were comfortable with the latest technologies.

C2 used ICT mostly for information search. She started with her academic needs when looking for exam sites, model question papers etc. ICT made it easy for her to download songs. 'Google search' makes it easier for her to do quick searches for anything and everything she needed even when visiting known sites.

The above findings give a different picture of women than in the literature. Women have overcome fear of technology and no longer have negative stereotypes about technology use. C1 and C2 both were confident when it came to ICT use for their needs. However, C1 seemed ahead of C2 in the use of multiple gadgets. C1 had many ICT devices which greatly facilitated her ICT acceptance and use. As a result, C1's EE (effort expectancy) was really low. She was already well versed with ICT features and use and her acceptance and use of ICT is high. This was evident from the interview and usage results.

C2 stood out form C1 in her need for relying on health information with ICT. She has used ICT for searching cancer-related information for her mother-in-law and also for her son's milk protein allergy. She comments-

"The boy has some milk allergy. I have Googled a lot about it. It is a rare, cow milk protein allergy. I have looked up a lot of sites. The last time I looked was one year back, as I don't get much time nowadays. This is not that common and we don't have many allergy specialists here. So, even pediatricians don't know much about this."

The above finding confirmed with the literature that women are active health information seekers and they employed weak search strategies as indicated by C2's search only strategy for health related information. However, she indicates– "*[If that's a medical problem] I refer to my textbooks. If I am still in doubt, I go to the internet or I ask some specialist."*

ICT had empowered both women to make informed decisions. Both reviewed items before a purchase. C2 looked for reviews mostly for big purchases such as camera or car where as C1 looked for wider reviews such as books, gadgets etc. Both looked at hotel reviews. C1 stated regarding information seeking for hotel reviews over the internet-

"[Internet] is a very open thing, so there is more information for me, especially when I look for products and even when we're planning a holiday. I feel Trip Advisor is a fantastic way of information seeking because many times, if you look at a hotel's website it always says good things and has good photographs. Same is the case even with products. Many times there are products that I am interested in. When I see it on the web and I do the forum review, I see people saying that it is a bad product. They say what the pros and cons are."

These experiences contribute to the PE from the device. Women's confidence level regarding the use of technology has generally been perceived to be very low. Online repositories have enabled them to be well informed. An instance of changing role of women is well reflected in the following anecdote by C1 on the purchase of her

iPad cover. In her words-

"So first [before buying the cover], I checked online and figured out the few brands and models that I liked. Then I thought maybe I should also take a look at the [physical] store. I went to the store and while I was talking to the sales person there, I figured out he didn't know much about this. I told him that I wanted a cover for my old iPad but instead he gave me the new one and said this would fit. I wasn't carrying my device but I knew this (the new iPad cover) won't fit [my] iPad. Mine is the old one. In all the reviews I had read online [I knew] this was lighter and the dimension was smaller. But the sales person said I was wrong and that he was dealing with Apple products every day. I did not want to argue but I trusted [the reviews]. I didn't want to take a chance and buy what he had suggested. It turned out to be right. Google was right. So I would like to rely on [the Internet]. There are so many objective reviews that you'll get and it's not a salesperson talking. It is mostly users talking and so it makes sense."

In another instance, C1 regretted the purchase of a book that she bought in spite of the negative reviews that it got.

As for C2, though she is not as frequent in seeking information as C1, nevertheless her information needs have equally enabled her in terms of family care. She was in the process of looking up information on land to be purchased by the family. Other information that she looked up was for her 1-year old such as car seats and other baby products. She had also looked for reviews of books and movies.

Contrary to the past women now are more empowered with information accessed using ICT. The degree of trust on information obtained from reviews was higher. As a matter of fact the elders and the male members of the family now relied on these women for any information they need from the Internet. ICT has enabled instant access to reviews and increased the trust factor on information obtained from this medium so much so that relying on misleading information even from 'experts', as the sales person above, would no longer be the only option.

It was evident from the interview that C2 led a much busier life compared to C1 yet her need to get quick information was facilitated by her laptop rather than her smartphone, which she mostly used for calls and messages. Her busy life may have made her EE high and the facilitating conditions were lower as compared to C1. Lack of time may have also led C2 to underutilize ICT as acceptance of a new technology needs time.

On the other hand C1 had to be in touch with her business even when out of office. She carried multiple ICT devices with her – smart phone, iPad and iPod. She accessed most of her information through her iPad and spent significant time browsing the internet throughout the day as a result of which over time she started exploring better and newer technology. For example she started printing and scanning using a new app on her iPad. For C2, she had expressed a desire to be more tech-savvy, however she felt constrained by her lack of time and resources. She thinks she would be able to get more done (in spite of her busy schedule) if she upgraded her connection to 3G.

This study's strength lay in looking at the changing IB of Indian women as brought about by ICT and their ICT adoption and use from the "user" perspective rather than the generalized work context.

5 Conclusion

This paper intends to study ICT adoption and use by Indian women from the two perspectives of Information behavior and technology adoption. The ICT usage patterns for both cases reaffirmed the role of factors such as social influence, performance expectancy and effort expectancy on predicting technology acceptance and use. The information behavior pattern for both cases displayed ICT as a facilitator, with C1 exhibiting high technology trust and skilled use as compared to C2.

Though the case studies provide scope for in-depth analysis, it is not generalizable, the findings can be used to design studies (qualitative/quantitative) that focus on specific aspects of women's information behavior, technology adoption or both in the Indian context. Studies can also focus on need for evolving ICT in order to facilitate technology adoption by women. The role of ICT in reshaping the image of women in the Indian society is also of paramount significance from a sociological perspective.

References

[1] D. Howcroft and E. M. Trauth, "The implications of a critical agenda in gender and

IS research," *Information Systems Journal*, vol. 18, pp. 185-202, 2008.

[2] C. Urquhart and A. Yeoman, "Information behaviour of women: theoretical perspectives on gender," *Journal of Documentation*, vol. 66 pp. 113 - 139, 2010.

[3] S. Jain, "ICTs and Women's Empowerment: Some Case Studies from India," Department of
Economics at LakshmiBai College, Delhi University. 2006.

[4] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, vol. 27, pp. 425-478, September 2003.

[5] V. Venkatesh, M. G. Morris, and P. L. Ackerman, "A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes," *Organizational Behavior and Human Decision Processes*, vol. 83, pp. 33-60, September 2000.

[6] H.-Y. Wang and S.-H. Wang, "User acceptance of mobile internet based on the Unified Theory of Acceptance and Use of Technology: Investigating the determinants and gender differences," *Social Behavior and Personality: an international journal*, vol. 38, pp. 415-426, 2010.

[7] M. L. Best and S. G. Maier, "Gender, Culture and ICT Use in Rural South India," *Gender, Technology and Development*, vol. 11, pp. 137-155, May 2007.

[8] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly*, vol. 36, pp. 157-178, March 2012.

[9] S. Ghobadi and S. Smith, "The Employment Of Online Communities Of Practice For Manifesting Feminist Behaviors Among Eastern Women," *PACIS 2011 Proceedings*, 2011.

[10] H. H. Gillard, Debra; Mitev, Natalie & Richaedson, Helen, "“Missing Women”: Gender, ICTs, and the Shaping of the Global Economy," *Information Technology for Development*, vol. 14, pp. 262-279, 2008.

[11] V. Nath, "Empowerment and Governance through Information and Communication Technologies:Women's Perspective," *International Information & Library Review*, vol. 33, pp. 317-339, 2001.

[12] D. Warner and J. D. Procaccino, "Toward Wellness: Women Seeking Health Information," *Journal of the American Society for Information Science and Technology*, vol. 55, pp. 709-730, 2004.

[13] J. Dunne, E., "Information seeking and use by battered women: A ‘‘person-in-progressive-situations’’ approach," *Library & Information Science Research*, vol. 24, pp. 343-355, 2002.

[14] L. Kvasny, E. Trauth, and A. Morgan, J., "Power relations in IT education and work: the intersectionality of gender, race, and class," *Journal of Information, Communication and Ethics in Society*, vol. 7, pp. 96-118, 2009.

[15] S. Sarker, S. Sarker, J. Ondrus, D. Jana, and M. Ahuja, "The Role Of Individual, Family-Related, And Organizational Factors In Shaping WLC In Offshoring Contexts: A Study Of European And Indian IT Professionals.," *PACIS 2011 Proceedings. Paper 162.*, 2011.

[16] S. Bardzell and E. F. Churchill, "IwC Special Issue "Feminism and HCI: New Perspectives." Special Issue Editors' Introduction," *Interacting with Computers*, vol. 23, pp. iii-xi, 2011.

[17] J. L. Quesenberry, E. M. Trauth, and A. Morgan, J., "Understnading the "Mommy Tracks": A Framework for Analyzing Work-Family Balance in the IT Workforce," *Information Resources Management Journal*, vol. 19, pp. 37-53, 2006.

[18] S. Halder, A. Ray, and P. K. Chakrabarty, "Gender differences in information seeking behavior in three universities in West Bengal, India," *The International Information & Library Review*, vol. 42, pp. 242-251, 2010.

[19] R. K. Yin, *Case Study Research: Design and Methods*, 3 ed.: Sage Publications, 2002.

Survey of fall detection techniques based on computer vision

Nuth Otanasap and Poonpong Boonbrahm
School of Informatics, Walailak University, Nakornsi Thammarat, Thailand
auisuke@gmail.com, poonpong@gmail.com

Abstract

Falls affect tens of millions of elderly people throughout the world, approximately 28-35% of people over the age of 65 falls each year. That means falls are the primary cause of injury related to death for elderly and the second leading cause of injury related death and set of symptoms after fall. Falling, the rapid movement from standing or sitting to lying position can be classified in to four phases as pre-fall, critical fall, post-fall and recovery phases. In this paper, we have reviewed and compared different fall detection techniques based on computer vision in four phases of fall. Examples of techniques founded are Bayesian network, Centroid, Eigen-space, Gaussian mixture model, Hidden Markov model, Neural network, Support vector machine and Velocity. Most of techniques are focused on critical phase and post-fall phase which mean that the falling person had been fallen already and may be injured. Some new sensor technologies included multi-camera systems were used to improve the results in pre-fall and early critical fall phase by combining information from several cameras and sensors to make decision and determine the direction of the signal from multiple sources that may track many people at the same time.

Keywords: Fall detection; Fall activities

1. Introduction

Statistics in Global brief for World Health Day 2012 [26] show that approximately 28-42% of elderly people fall each year. That means falls are the primary reason of injury related death for seniors and the second leading cause of symptoms after fall, which includes loss of autonomy, confusion, increased dependence, immobilization and depression. The fall therefore raises the interest of researchers, particularly preventive detection of the fall because there are inappropriately defined processes and can be approached using various methods. To reduce the risk of falling, the early detection of fall in pre-fall or critical fall phase therefore advances the interest of researchers. Several of research has been done in this area to develop algorithms and systems of preventive fall detection. The goals of this study were to classify the computer vision based approaches to prevented fall and to comparing the results of these studies.

2. Fall detection definition

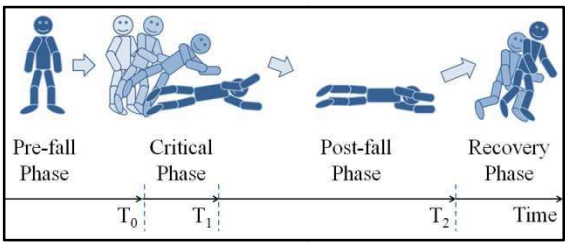


Fig. 1: Four phases of a fall event. [15, 19]

Firstly, as Fig. 1 the pre-fall phase is sudden movements directed to the ground as sitting or crouching down occasionally. The second phase is critical phase that very short period and most important of fall event. This phase can be detected by the body movement toward on the ground or by the impact shock with the floor. Next phase is the post fall phase which can be detected by motionless of human body after lying on the ground, lying position or by a motion absence significantly. Last phase is the recovery phase that faller get up from falling dawn. Explicitly, fall event in human can be described as the rapid movement from standing or sitting to lying position. During the critical fall phase there is free fall temporal time that continuous vertical speed increasingly.

3. Classification of fall detection techniques

As Fig. 2 [14] wearable device based, vision based and ambience sensor based are the three categories of existing fall detection methods as a hierarchy. Most of wearable sensors are based on accelerometers or based on gyroscopes which track the body movement automatically, not need intervention by human.

However, these kinds of sensors are often embarrassing to wear, and require batteries which need to be replaced or recharged frequently, that are important hindrance of these technologies.

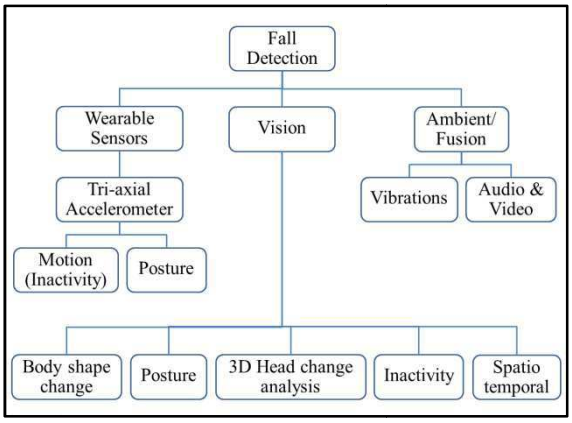


Fig. 2 Classification of techniques [14]

According to vision based device there are several techniques that are useful for detecting falling down of elderly people. As the following details, we have reviewed and compared different human motion tracking based on computer vision for understanding of fall detection techniques.

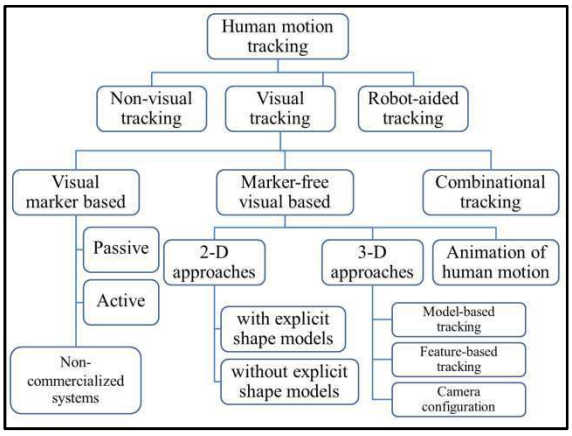


Fig. 3 Classification of motion tracking [29]

As Fig.3 [29], a tracking system can be non-visual, visual based or a combination of both. Visual tracking systems can be classified as either visual marker based or marker-free visual based as following details. Firstly, visual marker based tracking is a technique where cameras are adapted to track human movements, with identifiers position upon the human body. Consequently, each body part controls an unpredictable and complex motion trajectory, which may lead to contradictory and unreliable motion estimation. Moreover, disordered scenes or various lighting most likely divert visual attention from the real position of a marker. Visual marker based tracking is preferred to these kind of conditions that are frequently used as a human motion analysis standard due to their correct position information. One weakness of marker based using optical sensors is unable to detect joint rotation or body part overlapping and unable to render 3-D [21]. Secondly, marker-free visual based tracking systems utilize optical sensors (camera) that can produce continuous high resolution image to detect movements of body. Currently cameras are popularly used in surveillance applications with low cost and flexible to configure by user. However this technique requires complicated computation to error reduction, 3-D rendering, decreasing of data latency [4] and high speed camera [2] for high bandwidth. Lastly, Combination tracking systems take advantage of marker based and marker-free based technologies to reduce errors coming from using individual technology [23].

4. Related works

Generally, several kinds of techniques were used to detect falling down event in pre-fall, critical-fall or post fall phases. Some of them can only been used in one phase but some were used in several phases. Following details are briefly example of techniques that were used.

Cucchiara et al. [5] and Rimminen et al. [17] proposed a human posture classifier using Bayesian technique to detect falling in critical through post-fall phases. Both use shape, size and magnitude of the patterns or postures as an input for classification. Furthermore, centroids [8, 9, 10] are an effective way to measure object motion in space and can be computed as temporal averages of series of events. Researchers use a temporal moderate of the centroid motion events as an input to track fall risks and evaluate its dynamics in critical and post-fall phase.

An efficient method for activity recognition based on fusion of integrated spatio-temporal motion images and eigen-space techniques, mainly dedicated to fall detection in critical fall phase was proposed by Foroughi et al. [7] and Olivieri et al. [16]. The technique is used for accurate classification of motions and definition of a fall event in critical fall phase.

Silhouettes from multiple cameras [1] are used to build a 3-D estimation of the human by extracting feature from voxel person and used together with fuzzy inference to specify the state of fall in critical through post-fall phases. The resulting fuzzy rule base outputs are used to generate temporal linguistic summarizations and then temporally processed.

An automatically learning context-specific spatial model [12] in terms of semantic regions, particularly inactivity zones and entry zones, uses maximum a posteriori assessment of Gaussian mixtures (GM). Results of contextual model are presented using overhead camera sequences tracked using a particle filter that can be produced and unusual inactivity in post-fall phase to be detected. Rougier et al. [19] also presented a GM model (GMM) classification method to detect falls in critical through post-fall phases by analyzing human shape contortion during a video sequence. The silhouette of human along the video sequence was tracked by using a shape matching technique. The peak of fall is an important feature to indicate a fall in critical phase, but the lack of significant movement after the fall is also important for robustness when obstructions occur.

Hierarchical Hidden Markov Model (HHMM) and 3D human postures was proposed by Thome and Miguet [24] whose first layer states are related to the orientation of the tracked person in critical and post fall phase. They also proposed a multi-view method to achieve automatic detection of a falling person in video sequences, where motion is modeled using a layered hidden Markov model (LHMM) [25]. The posture classification is accomplished by a fusion unit that integrates the decision provided by the independently processing cameras in a fuzzy logic context.

The development of a vision system to detect natural events in a low-resolution image stream [3] concerns the assessment of algorithmic design decisions to increase detection reliability. As the approach of calculating vertical velocities from the infrared images and using a neural network to classify fall or non-fall in critical through post fall phase is sufficient to produce an actually fall detector. A comparison of the basic physics of falls to the ground that event can be explained in terms of the distributions of vertical velocities they generate, and the performance of the nearest neighbor classifier, an honest classifier.

A MEMS-based human airbag system [20] detects the complicated human motions and the recognition of a falling down motion in critical phase, which can be used to release an airbags. The system records human motion information using a high-speed camera then set up a human motion database that includes falls and normal motions, and use a support vector machine training process to classify. Moreover, a network of overlapping smart cameras [27] uses a decentralized process for computing inter-image homographies by calibrating only one camera. To evaluate fall detection procedure in post fall phase by both the Cyclops' algorithm and support vector machine (SVM) classifier, it uses a 3rd degree polynomial kernel along with three features which are extracted from the Cyclops' images.

The study of identify unique features of the velocity profile during normal and abnormal activities [28] made the automatic detection of falls during the critical phase of fall. Additionally the horizontal and vertical velocities of the trunk during normal activities [11] are within a well controlled distances, when the velocity in one direction increase, the velocity in the other direction normally did not. These two velocity characteristics could be used to differentiate fall movements from normal activities during the critical phase of the fall.

As the new hardware technologies and in particular digital cameras are now affordable and popularly used as tools for automatically assuring the human safety. Omni camera [13] and multi camera [6] vision system are proposed for detecting and tracking human and recognizing risk behaviors and events using posture based technique. Some posture methods to detect falls are based on a combination of motion history, human shape variation, event-inference, human aspect ratio and effective area ratio [11, 18, 22] that algorithm provides promising results on video sequences and simulated falls which can effectively prevent misjudgments and greatly increase the accuracy of detection results.

5. Summarization of phase detection

As Table 1, summarization of phase detection using several techniques has been presented.

There are 18 items of critical phase detection and 15 items of post-fall phase detection which several techniques. Most of them detect falling in critical and post fall phases sequentially that means pre-fall and early critical fall phase should be considered.

Table 1: Classification of Phase Detection

Detection Techniques	Critical Phase	Post-fall Phase	Input
Bayesian	[5, 17]	[5, 17]	posture
Centroid	[8, 9]	[9, 10]	centroid motion
Eigen-space	[7, 16]	-	saptio-temporal
Fuzzy	[1]	[1]	voxel person
GMM	[19]	[12, 19]	human motion
HMM	[25, 26]	[25, 26]	posture
Neural Network	[3]	[3]	velocity
SVM	[21]	[29]	human motion
Velocity	[11, 30]	-	human motion
Combination	[6, 11, 13, 18, 23]	[6, 11, 13, 18, 23]	posture

6. Conclusion and future work

We have reviewed different fall detection techniques as Bayesian, Centroid, Context Aware, Eigen-space, Fuzzy Logic, GMM, HMM, Neural Network, Posture, SVM, and Velocity. Several fall detection techniques are good solution for fall detection with high rates but mostly detect on lately critical phase or post-fall phase that mean fall person had been injured already. To reduce the risk of falling, the early detection of fall in pre-fall or critical fall phase therefore advances the interest of researchers. Some new sensor technologies could also be used to improve the recognition results in pre fall and early critical fall phase by combining information from several cameras and sensors by time synchronization like PrimeSence, Kinect or Xtion sensors that can provide color, depth, audio and video stream. Moreover, it features a microphone array that is possible to determine the direction of the audio source and optimizes skeleton tracking to recognize users as six people can be detected and one or two human can be tracked at one time.

References

[1] Anderson, D., Luke, R. H., Keller, J. M., Skubic, M., Rantz, M., & Aud, M. (2009). Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*, *113*(1), 80-89.

[2] Bhatnagar, D. K. (1993). Position trackers for Head Mounted Display systems: A survey. *University of North Carolina, Chapel Hill TR93*, *10*.

[3] Bromiley, P. A., Courtney, P., & Thacker, N. A. (2002). Design of a visual system for detecting natural events by the use of an independent visual estimate: a human fall detector. *Empirical Evaluation Methods in Computer Vision*, *50*.

[4] Bryson, S. (1996). Virtual reality in scientific visualization. *Communications of the ACM*, *39*(5), 62-71.

[5] Cucchiara, R., Grana, C., Prati, A., & Vezzani, R. (2005). Probabilistic posture classification for human-behavior analysis. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *35*(1), 42-54.

[6] Cucchiara, R., Prati, A., & Vezzani, R. (2007). A multi-camera vision system for fall detection and alarm generation. *Expert Systems*, *24*(5), 334-345.

[7] Foroughi, H., Naseri, A., Saberi, A., & Yazdi, H. S. (2008, October). An eigenspace-based approach for human fall detection using integrated time motion image and neural network. In *Signal Processing, 2008. ICSP 2008. 9th International Conference on* (pp. 1499-1503). IEEE.

[8] Fu, Z., Culurciello, E., Lichtsteiner, P., & Delbruck, T. (2008, May). Fall detection using an address-event temporal contrast vision sensor. In *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on* (pp. 424-427). IEEE.

[9] Leone, A., Diraco, G., & Siciliano, P. (2011). Detecting falls with 3D range camera in ambient assisted living applications: A preliminary study. *Medical engineering & physics*, *33*(6), 770-781.

[10] Lin, C. W., & Ling, Z. H. (2007, August). Automatic fall incident detection in compressed video for intelligent homecare. In *Computer Communications and*

Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on (pp. 1172-1177). IEEE.

[11] Liu, H., & Zuo, C. (2012). An Improved Algorithm of Automatic Fall Detection. *AASRI Procedia*, *1*, 353-358.

[12] McKenna, S. J., & Charif, H. N. (2004). Summarising contextual activity and detecting unusual inactivity in a supportive home environment. *Pattern Analysis and Applications*, *7*(4), 386-401.

[13] Miaou, S. G., Sung, P. H., & Huang, C. Y. (2006, April). A customized human fall detection system using omni-camera images and personal information. In *Distributed Diagnosis and Home Healthcare, 2006. D2H2. 1st Transdisciplinary Conference on* (pp. 39-42). IEEE.

[14] Mubashir, M., Shao, L., & Seed, L. (2013). A survey on fall detection: Principles and approaches. *Neurocomputing*, *100*, 144-152.

[15] Noury, N., Rumeau, P., Bourke, A. K., ÓLaighin, G., & Lundy, J. E. (2008). A proposal for the classification and evaluation of fall detectors. *Irbm*, *29*(6), 340-349.

[16] Olivieri, D. N., Gómez Conde, I., & Vila Sobrino, X. A. (2012). Eigenspace-based fall detection and activity recognition from motion templates and machine learning. *Expert Systems with Applications*, *39*(5), 5935-5945.

[17] Rimminen, H., Lindström, J., Linnavuo, M., & Sepponen, R. (2010). Detection of falls among the elderly by a floor sensor using the electric near field. *Information Technology in Biomedicine, IEEE Transactions on*, *14*(6), 1475-1476.

[18] Rougier, C., Meunier, J., St-Arnaud, A., & Rousseau, J. (2007, May). Fall detection from human shape and motion history using video surveillance. In *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on* (Vol. 2, pp. 875-880). IEEE.

[19] Rougier, C., Meunier, J., St-Arnaud, A., & Rousseau, J. (2011). Robust video surveillance for fall detection based on human shape deformation. *Circuits and Systems for Video Technology, IEEE Transactions on*, *21*(5), 611-622.

[20] Shi, G., Chan, C. S., Li, W. J., Leung, K. S., Zou, Y., & Jin, Y. (2009). Mobile human airbag system for fall protection using MEMS sensors and embedded SVM classifier. *Sensors Journal, IEEE*, *9*(5), 495-503.

[21] Sturman, D. J., & Zeltzer, D. (1994). A survey of glove-based input. *Computer Graphics and Applications, IEEE*, *14*(1), 30-39.

[22] Tao, J., Turjo, M., Wong, M. F., Wang, M., & Tan, Y. P. (2005, December). Fall incidents detection for intelligent video surveillance. In *Information, Communications and Signal Processing, 2005 Fifth International Conference on* (pp. 1590-1594). IEEE.

[23] Tao, Y., & Hu, H. (2003, September). Building a visual tracking system for home-based rehabilitation. In *Proc. of the 9th Chinese Automation and Computing Society Conf. In the UK*.

[24] Thome, N., & Miguet, S. (2006, December). A HHMM-based approach for robust fall detection. In *Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on* (pp. 1-8). IEEE.

[25] Thome, N., Miguet, S., & Ambellouis, S. (2008). A real-time, multiview fall detection system: A LHMM-based approach. *Circuits and Systems for Video Technology, IEEE Transactions on*, *18*(11), 1522-1532.

[26] World Health Organization. (2012). Good health adds life to years. Global brief for World Health Day 2012. *Geneva, Switzerland: World Health Organization. Accessed July*, *1*, 2012.

[27] Williams, A., Ganesan, D., & Hanson, A. (2007, September). Aging in place: fall detection and localization in a distributed smart camera network. In *Proceedings of the 15th international conference on Multimedia* (pp. 892-901). ACM.

[28] Wu, G. (2000). Distinguishing fall activities from normal activities by velocity characteristics. *Journal of biomechanics*, *33*(11), 1497-1500.

[29] Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control*, *3*(1), 1-18.

SVM Approach for Human Gender Recognition from Video Sequences

Ken Nishio† **Masao Izumi†**

†Graduated School of Engineering
Osaka Prefecture University, Japan

Abstract

Recently many researchers are trying to establish a human gender recognizer from images or videos. And it is commonly thought that facial images are useful for human gender recognition. But there still exists one of very challenging trials to recognize human gender from video or images in which whole bodies of human are captured and their faces are not so clear. In this article, we try to apply SVM to recognize human gender from video sequences. Results of our experiments show potential ability of our proposed method.

Keywords: Gender Recognition, HOG, SVM

1 Introduction

Nowadays there are a lot of surveillance cameras in cities for crime prevention. However these cameras can give us other kinds of information, such as marketing to know a flow of customers, what the customers tend to see, and so on [1]. Therefore, it has lately attracted considerable attention that we know human attributes. In some places, a face recognition system is used on a vending machine at the moment. The current understanding is that facial images are useful for human gender recognition and classification [2]. Though, it is not always able to detect faces in all situations. To better recognition without the facial images, here we study the way of using human whole body images. A few studies using the whole body images for gender recognition have been proposed [3], yet the study has a restriction on what human puts on, such as, females should dress skirts. As it can be noticed for the reason mentioned above, we think it can be valuable information to recognize human gender. Hence we try to recognize and classify human genders from non-facial images of, such as, surveillance cameras.

2 Proposed Methods

2.1 Outline

To recognize human gender, we should detect humans from videos, and divide them into two classes, male and female. Firstly, there is a human detection part using HOG feature. Secondly, obtained human images which we have classified as male or female by hand are applied as learning data to SVM for supervised learning. Finally, obtained human images which are not used as learning data can be classified by the learned SVM.

2.2 HOG-Based human detection

The human detection which is based on HOG feature is composed in OpenCV as samples [4].

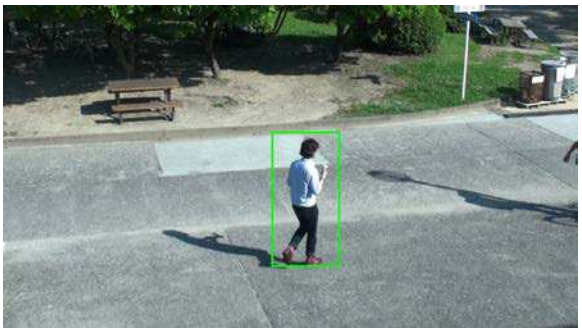


Figure 1. Human detection image

By using this OpenCV's HOG human detector, many images are found as human. However, there are a lot of other objects in them as noise. After removing them by hand, we divide human images into two, male class and female class.

2.3 SVM

Support Vector Machine (SVM) [5] is one of powerful two class pattern recognition discriminators. There are two classes; we regard female images as positive and male images as negative.

Then these are learned by SVM as a supervised learning. After this learning step, we put some images, which are not used in learning part, into the learned SVM discriminator.

Figure 2. Positive images (Females)

Figure 3. Negative images (Males)

2.4 Feature Vectors for SVM

It is a clearly one of key points what kind of feature vectors to be used for SVM in order to establish high recognition rate. This work is our first trial of gender recognition by using non-facial images, so we try to examine how effective it is to use luminosity information for gender recognition. Then we decided that we use greyscale luminosity of human bodies directly as feature vectors for SVM. The feature vectors are composed by luminosity values of each pixel of the images. The values are converted as greyscale data from color image data which have been detected by HOG human detector. These vectors are n-vector, this 'n' is decided by the size of images. As described in the following section, we use 100x200 resolution images as human body images, so in our experiments 'n' is 20000. We can divide the images into two classes by using the SVM which is learned by using learning feature vectors.

3 Results and Discussion

We have done the experiments of our proposed method with 3-Cross-validation using SVM. In 3-Cross-validation, all data are divided into three groups randomly, and two of three groups are used as learning data and the rest one group is used as test data. All combinations of selecting two groups as learning data are tested, then we finally get the results as the averages of all combinations.

3.1 Experimental condition

We try to discriminate between male and female, and divide 30 females (159 images) into 3 groups. Each group has 10 females as positive images. Then, male images (40 images) are prepared as negative images. When negative images evaluated as positive, we count the number of failures as False-Positive (FP).

The following shows the condition of this experiment.

- Resolution of camera images: 1440 x1080
- Positive and negative images (resized): 100x 200
- Positive images: 159
- Negative images: 40
- The number of females: 30
- The number of males: 8

3.2 Experimental results

Table 1 shows the results of 3-Cross-Validation. All groups' True-Positive rate is 100%. On the other hand, False-Positive rate are 30%, 20% and 27.5%. So the average of FP rate is about 25.8%.

Table 1. The result of 3-Cross-Validation

Group (images)	True-Positive (%)	False-Positive (%)
Group 1 (47)	47 (100%)	12 (30%)
Group 2 (62)	62 (100%)	8 (20%)
Group 3 (50)	50 (100%)	11 (27.5%)
averages	100%	25.8%

3.3 Discussion

In the results of our experiments, all 100% females and their images in 3 groups are identified as females. This shows that our trial could have an ability of the detection females from human images. However, this detector also identifies some male images as females in about 26%. That is because positive images are not many or some male's cloth resemble female's one. If the problem is caused by the former, we can avoid this to give more positive images to SVM. But in the latter case, we should find certain other features for more accurate classification. In future, we try to solve this problem. And we will get more samples and consider other features about difference of human gender.

4 Conclusion

In this research, we have tried to apply SVM approach to human gender recognition by using non-facial human body images. And the experimental results show the ability of human gender recognition without using facial data, just only using luminosity data of human body images. This is the first step for us to establish the method of human gender recognition by using non-facial images.

References

[1] Hitachi Solutions, Ltd. 2010-2013. http://www.hitachi-solutions.co.jp/company/press/news/2013/0304.html. (2013.5.24 accessible)
[2] Sony Corporation. http://www.sony.co.jp/SonyInfo/technology/technology/theme/sface_01.html. (2013.5.24 accessible)
[3] Hiroyuki Yamase, Kousuke Sato, Kunihiro Chihara, "Visual Discrimination of Gender by Standing Figure in Color Image", IEICE general conference, 1996. (in Japanese)
[4] OpenCV. URL:http://opencv.jp, (2013.5.24 accessible)
[5] SVM. http://docs.opencv.org/index.html, (2013.5.24 accessible)

Sentiment Polarity Estimation Using Term Chunking and Numeric Phrases in Fact Sentence

Hidekazu Yanagimoto†
†School of Engineering
Osaka Prefecture University, Osaka, Japan, 599-8531
hidekazu@cs.osakafu-u.ac.jp

Abstract

In the Internet vast amount of information is created and distributed to users. Hence, many users want how to understand them at glance and one of solution is sentiment analysis. I propose a sentiment polarity dictionary construction method using word chunking and numeric phrases. To construct a polarity dictionary with less human supports the proposed system uses semi-supervised learning algorithm which constructs the dictionary using word cooccurrence frequency. I construct two types of dictionaries; a numeric phrase dictionary and a chunked nouns and adjectives dictionary. In evaluational experiments Sentences including the numeric phrases are classify into a correct polarity group and their precision achieve about 88%. Combining method with a numeric phrase dictionary and a polarity dictionary improves 4% with respect to the precision.

Keywords: Natural Language Processing, Sentiment Analysis, Semi-supervised learning

1 Introduction

In the Internet vast amount of information is created and distributed to users. Hence, many users want how to understand them at glance and one of solution is sentiment analysis.

Sentiment analysis is a method that determine polarities of sentences using natural language processing. To determine a sentence polarity it uses a sentiment polarity dictionary which defines words and their sentiment polarities and count occurrence frequency of polarity-attached words in a sentence. The sentiment analysis consists of two phrase: sentiment polarity dictionary construction and polarity estimation using the dictionary. In a dictionary construction phase it is important to construct the polarity dictionary with less human efforts.

A polarity dictionary is constructed using existing dictionaries and thesaurus dictionaries. Takamura et al. [1] determined word sentiment polarities using a word network constructed with some dictionaries; WordNet[2] and other digitalized dictionaries. This approach can construct a large sentiment polarity dictionary.However, since it needs dictionaries to construct the network, it is difficult to use it to construct a sentiment polarity dictionary for specific domain, for example stock market news and so on.

A sentiment polarity dictionary is constructed using many unlabeled texts. Turney[3] assume that there often were positive words around a positive word and a word polarity was determined using a cooccurrence frequency. Hatzivassiloglou et al[4] assumed words with the same polarity was connected with appositive relation and a word polarity is determined using a conjunction. Nasukawa et al[5] estimated a word polarity using context coherence since sentiment polarity remains until contradictory conjunction happens. The approaches prepare a few polarity attached words (seed words), for example, ”good”, ”bad”, ”excellent” and so on, and estimate a polarity of an other word. Pan et al[6] focuses domain-specific words and domain-nonspecific words and developed a cross-domain sentiment classifier.

Next, we discuss sentiment analysis using a sentiment polarity dictionary. The main aim is to estimate polarity of sentences or documents including author’s opinion. However, sentences and documents are classified into two categories, Fact and Opinion[7]. Fact means a sentence describing what happens and needs interpretation to determine its polarity and Opinion includes author’s opinion obviously. It is more difficult to determine a polarity for Fact than for Opinion. Yu et al. [8] classified sentences into Fact sentences and Opinion sentences and estimated a polarity for each category. They found it is more

difficult to determine a polarity for a Fact sentence than for an Opinion sentence.

2 Proposed Method

2.1 Numeric Phrase in Fact Sentence

The numeric phrases means phrases that predicted a sentiment polarity. For example, "10 yens boost" and "5 points fall" are numeric phrases. The proposed method uses these phrase to estimate a sentiment polarity.

In the system I define some patterns extracted as a numeric phrase from stock market news. One of the patterns is a sequence of denomination and "boost" or "fall". The word sequence including "boost" has a positive polarity and the sequence including "fall" has a negative polarity.

Stock market news includes articles representing a price variation. In the case the news often uses a format of "**final price+variation**". Hence, I use a pattern like "**denomination+number**" as a positive phrase and "**denomination-number**" as a negative phrase.

2.2 Word Chunking

An ordinary sentiment analysis approach uses adjectives as a word including a sentiment polarity. Fact sentences include no adjectives and has no clue to estimate a polarity. Hence, to increase features the system uses nominalized verbs since such nouns include movement of price potentially. However, it is difficult to determine a polarity for such nouns since the meaning includes good side and bad side. For example, I consider "boost". Though "Stock price boost" means positive, "oil price boost" means negative. Hence, since nobody can determine a polarity using only "boost", I propose using a phrase chunked noun sequence as a feature.

2.3 Semi-Supervised Learning

The semi-supervised learning approach stars from a small dictionary where some words are registered according to specialists' knowledge. The approach expands the dictionary using rules that select words that are assigned with polarities. In a proposed approach we use a cooccurrence frequency with words that have positive or negative polarities.

The proposed approach uses a χ^2 statistic to evaluate the bias since a raw cooccurrence frequency is noisy. Equation (1) shows calculation of χ^2 statistic from cooccurrence frequencies.

$$\chi^2(w) = \sum_{w^s \in D_s} \frac{(\mathrm{freq}(w^s, w) - \frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N})^2}{\frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N}} \tag{1}$$

The $\mathrm{freq}(w_1, w_2)$ is cooccurrence frequency in a sentence and the $\mathrm{freq}(w)$ is occurrence frequency in a text corpus. The N denotes the number of all the sentences in the text corpus and the D_s denotes a set of seed words.

We use an improved χ^2 statistics to neglect difference between positive seed occurrence and negative seed occurrence.

$$\chi^2_P(w) = \frac{1}{|D^P_S|} \sum_{w^s \in D^P_s} \frac{(\mathrm{freq}(w^s, w) - \frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N})^2}{\frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N}} \tag{2}$$

$$\chi^2_N(w) = \frac{1}{|D^N_S|} \sum_{w^s \in D^N_s} \frac{(\mathrm{freq}(w^s, w) - \frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N})^2}{\frac{\mathrm{freq}(w^s)\mathrm{freq}(w)}{N}} \tag{3}$$

The D^P_S and the D^N_S are a set of words with positive polarities and a set of words with negative polarities respectively.

When the χ^2 static is over a threshold, which are defined previously, the word is registered in the dictionary.

$$w = \begin{cases} \text{positive} & (\chi^2_P(w) > \text{threshold}) \\ \text{negative} & (\chi^2_N(w) > \text{threshold}) \end{cases} \tag{4}$$

When a negation follows polarity-attached words, the polarity is opposite since negation works as a factor to reverse the polarity. Hence, counting concurrence frequency, my system considers negations.

2.4 Sentence Polarity Estimation

Since the dictionary consists of words and their polarities, a polarity of a sentence is determined according to how many positive or negative words are included. The above judgement uses a assumption that is that negative sentences include many negative words and positive sentences include many positive words. When there

Table 1. Dictionary size

Dictionary type	The number of words
denomination dictionary	44
word-based dictionary	183
phrase-based dictionary	376

are the same number of positive words and negative words in a sentence or there are no polarity-attached word in a sentence, the sentence is estimated as neutral one.

Positive words attached with a negation is regarded as negative words and negative words attached with negation as positive words.

3 Experiments

3.1 Data Set

In this experiments we used T&C news, which was one of stock market news delivery services, as training data for dictionary construction and test data for evaluation. T&C news is written in Japanese and consists of 62,478 articles in 2010. Previously 100 articles are evaluated by an expert manually and have a sentiment polarity. Moreover, 100 articles include 939 sentences and the sentences have a sentiment polarity evaluated by the expert manually, too. The remaining 62,378 articles consist of 431,223 sentences and a polarity for them are unknown.

I used 431,223 sentences to construct a sentiment polarity dictionary. I used adjectives and nominalized verbs in the sentences. To chunk nouns and nominalized verbs the proposed method check 5 precedent words.

In this experiment we set the threshold, which is in Equation (4), as 300.0. The larger threshold we set, the smaller dictionary is constructed.

3.2 Results

I show a sentiment polarity dictionaries using semi-supervised learning approach and hand-crafted numeric phrase dictionary in Table 1.

In Table 2 registered words in the phrase-base dictionary are sampled. Although ”boost” has a positive polarity originally, the polarity varies according to a chunked noun. For example, ”stock price boost” has the same polarity as ”boost”. On the other hand, ”material cost boost” and ”yen price boost” has the opposite polarity since they have a bad effect on the stock market. On the other hand, though ”profit increase” is a positive phrase, ”actual profit increase” become a negative phrase.

Table 2. Example of registered words in the phrase-base dictionary

Phrase	Polarity
boost	1
financial stock price boost	1
material cost boost	-1
yen price boost	-1
profit increase	1
actual profit increase	-1

Table 3. Sentiment polarity estimation using numeric phrases

		Estimation		
		p	n	-
Label	p	54	1	0
	n	1	17	0
	-	5	3	0

Using Table 3, I discuss a performance of sentiment analysis with numeric phrase. Sentences including the numeric phrases are 81 sentences in 939 sentences. The precision achieved about 88%. This result denotes the dictionary includes appropriate phrases for polarity estimation.

Table 4 and Table 5 show a classification result using only a word-base dictionary and using only a phrase-based dictionary respectively. The approach with word-based dictionary attached a positive label or a negative label to originally neutral sentences. Because of a word-based dictionary a registered word happens in a sentence more frequently than a phrase. Hence, the approach with a word-based dictionary overestimates polarities of sentences.

Table 6 and Table 7 shows results of combination approaches. In first step, the system determines sentence polarities using a numeric phrase dictionary. In second step, polarity unknown sentences are processed with a sentiment polarity dictionaries; a word-based dictionary and a phrase-based dictionary. Both of combination approaches are superior to non-combination approaches. Using numeric phrases increases about 4% precision. In Table 8 a combination approach with numeric phrase and word-based dictionary achieves the highest precision.

Table 4. Sentiment polarity estimation using a word-based dictionary with all sentences

		Estimation		
		p	n	-
Label	p	192	68	142
	n	76	98	89
	-	82	64	128

Table 5. Sentiment polarity estimation using a phrase-based dictionary with all sentences

		Estimation		
		p	n	-
Label	p	180	55	167
	n	73	95	95
	-	69	68	137

4 Conclusion

I proposed a sentiment analysis method using term chunking and numeric phrase in Fact sentences. I pick up price variation phrases as numeric phrase to determine sentiment polarities. In polarity dictionary construction phase, semi-supervised learning is used to decrease human efforts. Considering numeric phrase, the proposed method improves about 4% with respect to precision.

Acknowledgment

I would like to thank Centillion Company for giving us T&C news.

References

[1] H. Takamura, T. Inui, and M. Okumura. Extracting semantic orientations of words using spin model. *Proceedings of the 43th Annual Meeting of the Association for Computational Linguistics*, 2005.

[2] C. Fellbaum. *WordNet: An Electronic Lexical Database, Language, Speech, and Communication*. MIT Press, 1998.

[3] P. D. Turney. Thums up? thums down? semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 2002.

[4] V. Hatzivassiloglou and R. K. McKeown. Predicting the semantic orientation of adjectives. *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, 1997.

[5] T. Nasukawa and H. Kanayama. Acquisition of sentiment lexicon by using context coherence. *NL-162-16*, 2004.

[6] J. S. Pan, X. Ni, J. Sun, Q. Yang, and Z. CHen. Cross-domain sentiment classification via spectral feature alignment. *Proceedings of WWW2010*, 2010.

[7] R. Feldman. Techniques and appications for sentiment analysis. *Communication of the ACM*, 56, 2013.

[8] H. Yu and V. Hatzivassiloglou. Towards answering opinion questions: separating facts from opnions and identifying the polarity of opinion sentence. *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, 2003.

Table 6. Sentiment polarity estimation using a word-based dictionary with sentences not including numeric phrases

		Estimation		
		p	n	-
Label	p	166	58	123
	n	70	89	86
	-	73	63	130

Table 7. Sentiment polarity estimation using a phrase-based dictionary with sentences not including numeric phrases

		Estimation		
		p	n	-
Label	p	155	49	143
	n	68	87	90
	-	61	66	139

Table 8. Precision for polarity estimation

Dictionary	Precision
word-base	0.445
phrase-base	0.439
numeric phrase+word-base	0.486
numeric phrase+phrase-base	0.481

Text Classification Tool for Sentiment Analysis on Social Media Monitoring in Retail Business

Vasutan Tunbunheng **Datchakorn Tancharoen**
Faculty of Engineering and Technology
Panyapiwat Institute of Management
{vasutantun, datchakorntan}@pim.ac.th

Abstract

Currently, a social media plays an important role in the business. The customer opinions are an important factor to improve the services in the retail business. The opinions are broaden rapidly in the cyber world especially in social media. Therefore, the retail business is interested in the sentiment analysis to obtain the useful information which the customers are thinking about their products. However, the number of opinions are large and it is difficult to analyze the information manually. In order to analyze such a big data, the input opinion messages are grouped by using keywords. The sentiment analysis could be scoped down to the specific topics and analyzed by the text classification tool. The information is analyzed from the customer opinions in the social media. Then the negative opinions are discussed to find the solutions to improve the quality of services in the business.

Keywords: sentiment analysis, text classification, social media monitoring, retail business

1 Introduction

Social media becomes an important issue on the internet. People can share their experiences and communicate with others around the world using social media. The examples of social media include blogs, webboard, chat room, YouTube, Twitter, facebook as shown in Figure 1. Most of information exchanged between people is in the form of text data. Many businesses are interested in the sentiment analysis to obtain useful information in order to improve their products and plan for the future action for their businesses.

The text data on social media is usually written as same as people talking in daily life so the text data is used to analyze the customer opinions. However, it is difficult to analyze and understand the meaning of text data individually without classification. There were some researches on natural language that try to obtain the information from the text data such as Y. Duan [1] and M. Bates [2]. However, the information could be analyzed manually by the group of workers since the automated sentiment analysis sometimes fail in the business. Social network analysis and mining was also performed in business applications [3]. The text categorization and text mining were the necessary tools for information discovery [4], [5]. Several systems have been built to quantify opinion from product reviews. In [6], they identify which sentences in a review are the subjective characters to improve sentiment analysis. Local sentiment was identified to be more reliable than global document sentiment. They focus on identifying the orientation of sentiment expressions and determining the target of these sentiments [7]. The given topic using natural language processing technique was analyzed to extract the sentiments [8]. Therefore, the text classification is useful to support the sentiment analysis on the big data in social media. The background of the research for social media monitoring in the retail business is stated in section 2. The text classification tool is explained in section 3. The experiment of this project is shown in section 4 and section 5 is the conclusion.

Figure 1. Various social media on the internet

2 Social Media Monitoring

Social media is a new media which has the impact to the customers. Many businesses would like to collect the comments and messages of their customers from social media. Thus, we plan the platform for social media monitoring. This platform includes a robust crawler to collect the big data from social media, the large scale database for service APIs, and also the sentiment analyzer to look the customer opinions. There is a work in the retail business to monitor what the customers are talking about the company on a social media espeically on facebook. They collected the text opinions on the facebook including some keywords and classified the opinions to the group of positive, neutral and negative opinions as shown in Figure 2. The problem is that this work is difficult to do manually without the tool to analyze the text data since the number of messages is very large. After the business is monitoring the customer opinions, they found that the text data could be grouped into a certain number of topics. In addition, the text data on each topic always contain the same group of keywords. This paper presents the text classification tools using keywords to classify the text data into different groups as in the following section.

Figure 2. Opinion analysis and classification in the positive, neutral, and negative

3 Text Classification Tool

The text data and group of keywords, which stored in two different files are the input for a program named "Text Group" in the format of Microsoft Excel.

The sheet format of TEXT data is shown in Figure 3. It is consisting of messages. Each message is on each row without space row. Thus, the space row is marked as the ending of messages list.

The sheet format of GROUP data is shown as Figure 4. The first column is group name and the second column is the keywords belong to the group. If the message in TEXT file is consisting of the keyword, it is classified to the corresponding group name. The pair of keyword and group name is loaded until the space row is reached.

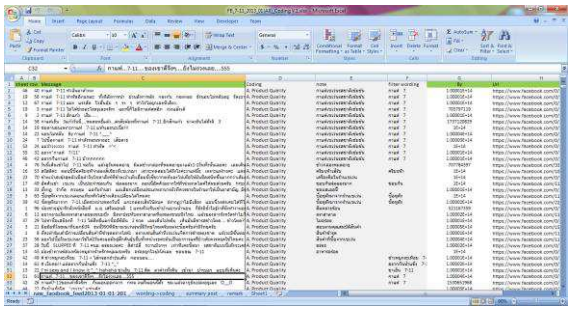


Figure 3. TEXT sheet format

	Coding	Word filter 1
2	H. Post by Store staff,	7-11 รวมทีม
3	H. Post by Store staff,	7-11รวมทีม
4	H. Post by Store staff,	part time ที่เชเว่น
5	H. Post by Store staff,	กลับปัยหาเพิลๆที่ 7-11
6	H. Post by Store staff,	คนทำงาน 7-11
7	H. Post by Store staff,	คิด ถึง เพื่อน ๆ ที่ 7-11
8	H. Post by Store staff,	คิดถึงทุกคน 7-11
9	H. Post by Store staff,	คิดถึงพี่ๆ น้องๆที่ร้าน 7-11
10	H. Post by Store staff,	คิดถึงเพื่อนๆพี่ๆน้องๆที่ 7-11
11	H. Post by Store staff,	ใครเคยทำงานเชเว่น
12	H. Post by Store staff,	ใครสนใจทำงานร้าน 7-11
13	H. Post by Store staff,	ชาว 7-11 ที่รัก
14	H. Post by Store staff,	ตรวจ 7-11
15	H. Post by Store staff,	ต้องอยู่ร้าน 7-11
16	H. Post by Store staff,	ทามงานที่ 7-11
17	H. Post by Store staff,	ทำงาน 7-11
18	H. Post by Store staff,	ทำงาน ที่ 7-11
19	H. Post by Store staff,	ทำงาน พาร์ทไทม์ 7-11
20	H. Post by Store staff,	ทำงาน7/11

Figure 4. GROUP sheet format

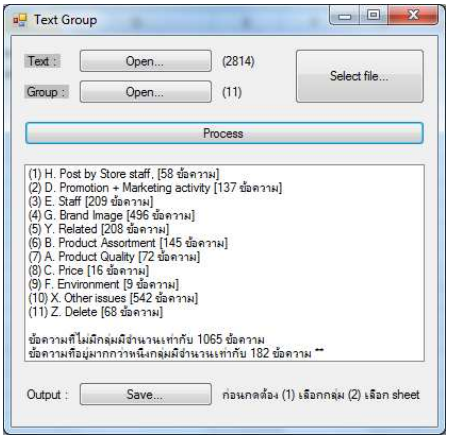


Figure 5. The classification results

When the text classification finished, every group name was listed under the "Process" button followed by the number of messages inside the group as shown in Figure 5. The number of messages which are included into more than one group and are not included in any group are shown below the group list.

After input text data and keywords, the "Process" button become enable and each message is included into group if it contains corresponding a keyword. In the case that the message belong to more than one group, the longer keyword is selected. The algorithm is shown as the flowchart in Figure 6.

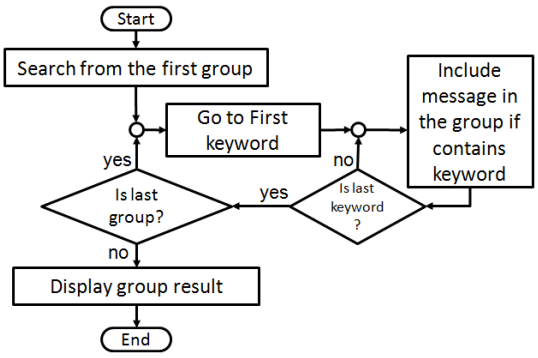


Figure 6. Flowchart to process

4 Experiments

In the experiments of text classification, the text data is obtained from social media monitoring service. It collects the post messages from social media named "facebook". The user can update keywords into the "Text Group" application and process to update the result repeatly. Thus, the number of messages belong to the unknown group is reduced everytime of the keyword updating. There are 11 groups in the current developping project. The text data is consisting of 2,814 messages from the web crawler in a month. The result can be shown as Table 1.

If the messages belong to the multiple groups, the longer keyword is selected. Then, the number of message belong to multiple groups is counted. It prompts for user to update the keyword again to reduce the number of candidate. On the other hand, the number of messages which are not belong to any group is also counted. The result can be shown as Table 2.

Table 1. Experiments of this project

Group Name	#Messages	Percent%
Post by Store Staff,	58	3%
Promotion Activity	137	7%
Staff	209	10.7%
Brand Image	496	25.3%
Related	208	10.6%
Product Assortment	145	7.4%
Product Quality	72	3.7%
Price	16	0.8%
Environment	9	0.5%
Other issues	542	27.7%
Delete	68	3.5%

Table 2. The number of special case messages

Special Case	#Messages	Percent%
Multiple groups	182	6.46%
Unknown group	1065	37.84%

Figure 7 presents the circle chart for the text group classification. The percentage of posts is shown in each message group. We found that the brand image and other issues are the topics that people usually talk about this retail business. In addition, this text classification tool also presents the visualization of message posts everyday during January 2013 as presented in Figure 8. We found that there were some topic issues which increased rapidly in the middle of the month. Thus, it is an issue that the business has to find out the solution to improve the process or gather the ideas for the business on this issue.

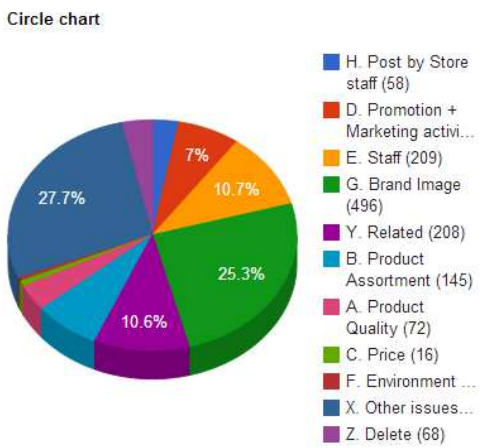


Figure 7. Text group classification experi-ments in the circle chart

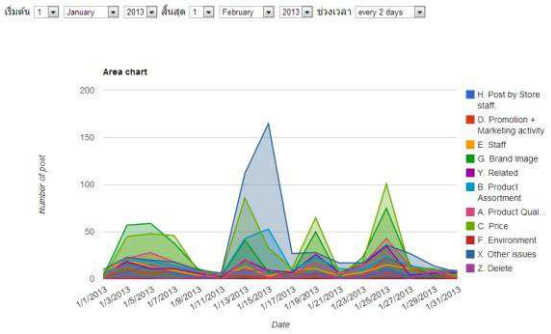


Figure 8. The number of posts in each group category during a month

5 Conclusion

This paper presented the social media monitoring platform and text classification tool for sentiment analysis. The sentiment analysis is a hard work since the number of messages is large. The messages can be classified by using keywords to separate the different groups and we can focus the messages on specific topic. It can reduce the time to analyze the large amount of text data. However, some messages cannot be classified. We need to repeat the process of text classification. The keywords are also updated manually to reduce the number of messages which included in many groups or not included in any group. This is a useful tool to classify and monitor the message contents on social media in retail business. This tool can show the visualization of group text messages in a period of time. This is helpful for the retail business to monitor the social media and scope into the specific group for sentiment analysis.

References

[1] Yucong Duan, Christophe Cruz, "Formalizing Semantic of Natural Language through Conceptualization from Existence", *International Journal of Innovation, Management and Technology*, Vol. 2, No.1, pp. 37-42, 2011.

[2] M. Bates, "Models of Natural Language Understanding". *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 92, No. 22, pp. 9977–9982, 1995.

[3] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes, "Social Network Analysis and Mining for Business Applications". *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 3, Article 22, April 2011.

[4] S. Shehata, F. Karray, and M. Kamel, "A Concept Based Model for Enhancing Text Categorization", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pp. 629-637, 2007.

[5] Ning Zhong, Yuefeng Li, and Sheng Tang Wu, "Effective Pattern Discovery for Text Mining". *IEEE Transactions on Knowledge and Data Engineering*, Vol.24, No.1, pp. 30-44, 2012.

[6] Pang, B., Lee, L., "A Sentimental Education: Sentiment Analysis using Subjectivity Summarization based on Minimum Cuts", *Proceedings of International Conference on ACL*, pp. 271-278, 2004.

[7] Nasukawa, T., Yi, J., "Sentiment Analysis: Capturing Favorability using Natural Language Processing", *The Second International Conferences on Knowledge Capture*, pp.70-77, 2003.

[8] J. Yi, T. Nasukawa, R.B., Niblack, W., "Sentiment Analyzer: Extracting Sentiments about a Given Topic using Natural Language Processing Techniques", 3*rd IEEE International Conference on Data Mining (ICDM)* ,pp.423-434, 2003.

A Novel Interface of Interactive Evolutionary Computation for Automatic Sentence Generator

Kiyohito Fukuda† **Naoki Mori**† **Keinosuke Matsumoto**†
†Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuencho, Nakaku, Sakai, Osaka 599-8531, Japan
fukuda@ss.cs.osakafu-u.ac.jp

Abstract

Recently, automatic sentence generation by the computer has become an important topic of natural language processing field. However, it is difficult for computer to generate human-like sentences because computer can not evaluate the quality of sentences. To introduce the human evaluation results into computers, IEC (Interactive Evolutionary Computation) has been proposed and reported lots of applications for design fields. In this paper, we proposed a novel interface of IEC for automatic sentence generation. To apply IEC for automatic sentence generation, 2-dimensional map type interface is helpful because human evaluates sentences by different type of features. We also created the GUI system of automatic sentences generation to confirm the effectiveness of proposed interface.

Keywords: Interactive evolutionary Computation, Mushroom Produce Atelier, 2-dimensional Map Type Interface

1 Introduction

Recently, automatic sentence generation by the computer has become an important topic of natural language processing field. It is becoming possible to generate short sentences or fix phrases automatically, but still difficult for computer to generate long story. It is also difficult to generate a short sentence with considering human preference because there is no criteria to evaluate sentence quantitatively. One of the most important point is that we feel differently about even seeing same sentence. In order to solve this problem, the novel evolutionary computation (EC) technique called *Interactive Evolutionary Computation* (IEC) [1] [2] [3] has been proposed and applied to fields such as design and music. IEC evaluates individuals of EC based on human evaluation. We use IEC for automatic sentence generation because IEC is powerful method to estimate human preference or sense. However, there exists a problem that users could not tried huge number of trials in selection phase of EC to reflect their preference.

In this paper, we proposed a novel interface of IEC for automatic sentence generator in order to estimate human preference in details. The proposed interface has two evaluation axis in 2-dimensional map. We can set the meaning of these axis such as positive/negative or formal/informal according to the target problem. We show the effectiveness of proposed interface by means of our automatic sentence generation system called “Namekoubou” (Mushroom Produce Atelier) [4].

We describe the outline of EC and IEC in section 2, and explain the outline of our proposed automatic explanation generator called Namekoubou in section 3. In section 4, we explain the setting of IEC. In section 5, we propose a novel interface of IEC, and show the experiment results in section 6. Finally, we conclude the paper in section 7.

2 Evolutionary Computation

Evolutionary Computations (ECs) are a search and optimization technique based on the mechanism of evolution. The *Genetic Algorithm* (GA) is one of the popular EC and have been researched in various fields. ECs have three main genetic operators called crossover, selection and mutation for search.

2.1 Interactive Evolutionary Computation

Interactive Evolutionary Computation (IEC) [1] [2] [3] is one of the interesting branch of EC. IEC decides individual’s fitness by human evaluation to solve the problem like music or CG art. IEC is expected to solve the problem with human emotion or sensitivity by user’s evaluations.

2.2 Applications of IEC

IEC is often applied to various fields such as design and music where evaluation strongly depends on user's emotion. There have been reported lots of researches of IEC such as CG composition [5], creation of portrait [6] and GenJam which is a system evolving jazz [7]. However, there is few research of applying IEC to problem of natural language processing field.

2.3 Problems of IEC

To apply IEC for problems of human preference or sense, human interface for IEC selection is essential to obtain good results. In this paper , we try to estimate multi-objective human preferences in order to reflect detail human preferences to IEC evolution process.

3 Sentence Generation

Recently, lots of study have been reported about automatic sentence generation. However, to create a long story by computer has been still quite difficult because computer is hard to maintain the logical consistency of background of the whole story. Therefore, we focus on the short sentence of explanation under the specific situation.

3.1 Mushroom Produce Atelier

We focused on the popular smartphone application named "Touch Detective Mushroom Garden" and have proposed *Mushroom Produce Atelier* (Japanese title Namekoubou) [4] which can create the name and explanation of a new mushroom character based on the user's input keyword. The produced explanation has only 1 sentence, and accepts surrealistic contents in order to reduce the difficulty of maintaining the logical consistency.

3.2 Outline

Figure 1 shows the outline of *Mushroom Produce Atelier*. User inputs a favorite noun to the system, then *Mushroom Produce Atelier* produces the name and explanation of new character related to input word automatically.

3.2.1 Human Evaluation

Mushroom Produce Atelier (MPA) makes several candidate sentences and user can select their favorite one. Current evaluation interface is the one slider bar to decide the fitness value of sentences. However, lots of user would like to input

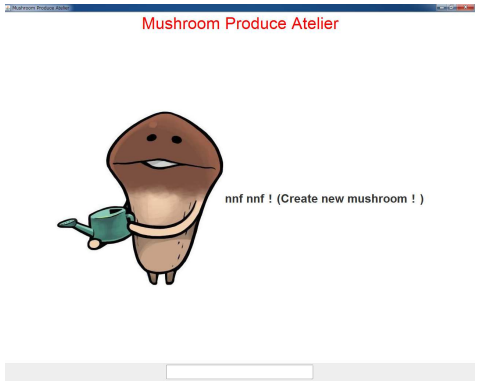


Figure 1. Outline of *Mushroom Produce Atelier*
©Beeworks/SUCCESS All rights reserved.

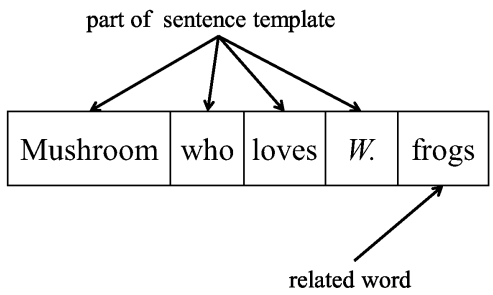


Figure 2. Genotype of Individual

not only fitness but another features in this selection process. This is the one of the biggest problem of MPA.

4 Setting of IEC

In this section, we show the setting of IEC in MPA.

4.1 Individual

In this study, IEC individual represents a sentence of certain mushroom. Each individual is created by combining the sentence template and specific word. Followings are the example of templates.

- Mushroom who loves W.
- Is he W? No... he is also mushroom...

,where W is a specific word decided by the relation to user's input keyword by means of the wikipedia thesaurus. We made 100 templates before simulation and assign randomly to all individuals in the initial population. Figure 2 shows the genotype of individual in this study.

4.2 Fitness

Proposed interface can represent 2-objective fitness value f_α and f_β are defined as follows.

$$f_\alpha = 0.2 f_y + 1.6 f_y f_x \tag{1}$$

$$f_\beta = 0.2 f_y + 1.6 f_y (1 - f_x) \tag{2}$$

Where, f_x means a evaluation value obtained from x axis and f_y means a evaluation value obtained from y axis, explained in section 5.

4.3 Genetic Operator

4.3.1 Mutation

In current setting, both related word W and sentence template are the target of mutation. Alleles of sentence template are 100 sentence templates made before simulation, and alleles of W are about 60 words from the related words of the input keyword by wikipedia thesaurus and wordnet. If we get $n < 60$ related words, the number of alleles is set to n.

4.3.2 Crossover

We create 2 individuals from 2 parents by one point crossover. Here, we represent 2 offsprings by A and B. In this study, we fixed the crosspoint to boundary between template and related word. We choose a only 1 offspring with probability $P(\mathrm{A})$ and $P(\mathrm{B})$ from A and B. $P(\mathrm{A})$ is the probability that individual A is selected in this procedure. $P(\mathrm{A})$ and $P(\mathrm{B})$ are defined as follows.

$$P(\mathrm{A}) = \frac{2|f_\mathrm{A} - f_\mathrm{B}| + 1}{2} \tag{3}$$

$$f_\mathrm{A} = \frac{f_{\mathrm{A}\alpha} + f_{\mathrm{A}\beta}}{2} \tag{4}$$

$$P(\mathrm{A}) + P(\mathrm{B}) = 1 \tag{5}$$

Where, $f_{\mathrm{A}i}$ means f_i of individual A.

4.3.3 Selection

We define L as population size. $L-2$ individuals of generation $t+1$ are created by mutation and crossover based on $2(L-2)$ individuals selected by roulette selection of generation t. An elite individual in generation t is always selected in generation $t+1$. Finally, 1 individual of generation $t+1$ is created randomly.

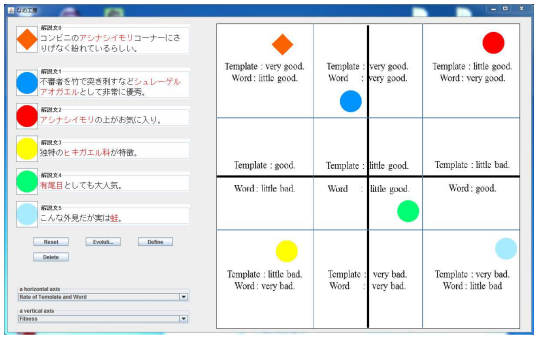


Figure 3. Outline of Proposed Interface

5 Proposed Interface

Evaluation of sentence is multi-objective problem. Some user may focus on words, and the other may select according to the template. The followings are the example of important objective functions of sentence.

- Total impression
- Sentiment of words
- Unpredictability

However, our previous system only consider the total fitness and can not evaluate other features of sentences. To solve this problem, we introduce 2-D map type interface for IEC. User only put a individual on adequate place to reflect their preference. User can easily change the type of axis. The result is feedback to setting of IEC genetic operators. Current version tried to evaluate words and templates separately, and result is reflected to parameters of mutation and crossover. There have been proposed several research about IEC interface such as making user model in computeror mapping all objective function space into 2-D plane. In this study, we proposed a novel interface of IEC for sentence generation. Figure 3 shows the example of proposed interface. We can evaluate each individual by arranging on 2-D map. The location of each icon on 2-D map represents fitness.

6 Computer Experiment

6.1 Experiment Method

To confirm the effectiveness of proposed method, computer simulation is carried out. Table 1 shows parameters of the IEC. We compared proposed method with previous version. Processes of the experiment is following:

Table 1. Parameters of the IEC

Population Size	6
Crossover Rate	1.0
Mutation Rate (word) (template)	 1.0 - f_α 1.0 - f_β
Selection Type	Roulette, Elitism
Number of Subjects	11

f_α and f_β are obtained from user's action on selection phase and both range are [0, 1].

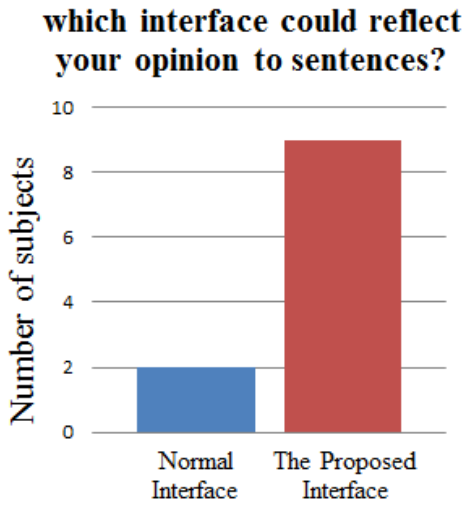


Figure 4. Experiment Results

1. A user tries MPA with previous interface to create a favorite sentence.
2. A user tries MPA with the proposed interface to create favorite sentence.
3. A user fills out a questionnaire about interactive process of MPA.

6.2 Experiment Results

Figure 4 shows the results of a questionnaire that "Which interface could reflect your opinion to sentences?". This result shows that the proposed interface can reflect user's opinion to sentences more than normal one. Therefore, we can search solution more efficiently than ever before.

7 Conclusion

In this paper, we proposed a novel 2-D map type IEC interface for automatic sentence generation system called *Mushroom Produce Atelier*. We showed the basic concept and GUI outline of the proposed interface. We explained that users can evaluate sentences subjectively by the proposed interface. Creating general sentences is important further study.

This work was supported by JSPS KAKENHI Grant, Grant-in-Aid for Scientific Research(C), 2500208, 2010-2013.

References

[1] Hideyuki Takagi. Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation. *Proceedings of the IEEE*, 89(9):1275-1296, 2001.

[2] Christopher L. Simons and Ian C. Parmee. EA cross-disciplinary technology transfer for search-based evolutionary computing: From engineering design to software engineering design. *Engineering Optimization*, 38(5):631-648, 2007.

[3] Shinya Henmi, Shino Iwashita, and Hideyuki Takagi. Interactive evolutionary computation with evaluation characteristics of multi-iec users. *IEEE Int, Conf, on Systems, Man, and Cybernetics (SMC2006)*, pages 3475-3480, 2006.

[4] Kiyohito Fukuda, Naoki Mori, Keinosuke Matsumoto: "Proposal of an Automatic Sentence Generator of under Specific Situation", The 57th Annual Conference of the Institute of Systems, Control and Information Engineers, 337-3, 2013.

[5] McCormack, J.: Interactive Evolution of L System Grammars for Computer Graphics Modelling, in Green, D. G. and Bossomaier, T. eds., Complex Systems: from Biology to Computation, IOS Press, Amsterdam, Netherlands, pp.118-130, 1993.

[6] Toshiaki Nakasu, Naiwala Pathirannehelage Chandrasiri, Takeshi Naemura and Hiroshi Harashima: "The NIGAO Interactive Facial Caricature Drawing System", The Institute of Image Information and Television Engineers, 61(6), 779-788, 2007.

[7] Biles, J. A.: GenJam: A Genetic Algorithm For Generating Jazz Solos, in *Int'l Computer Music Conf. (ICMC'94)*, Aarhus, Denmark, pp.131-137, 1994.

Robust Scheduling by Hybrid Method of Policy Tree and Estimation of Distribution Algorithm and Application to Resource Constrained Project Scheduling Problem with Rework Uncertainty

Jing Tian†* **Xinchang Hao†** **Tomohiro Murata†**

† Graduate School of Information, Production and Systems,
Waseda University, Kita-Kyushu, Japan
tasei@ruri.waseda.jp

Abstract

This paper presents using policy tree and estimation of distribution algorithms (EDAs) to solve resource-constrained project scheduling problem (RCPSP) with considering rework with probability. Firstly, the algorithm adopts policy tree to model the project scheduling problem, which contains decision variables, stochastic variables and precedence relations, trying to make proactive scheduling and reactive scheduling by structure of tree. Secondly, in order to avoid the memory problem associated with large tree, a linear function is used to generate policy tree, which is acted as the rule to generate dynamic scheduling for solution. Experiments results based on a set of benchmarks (PSPLIB) and comparisons with some existing methods demonstrate our approach with high effectiveness and tolerance with uncertainty.

Keywords: Proactive/reactive Scheduling, Uncertainty, Resource-Constrained Project Scheduling Problem, Policy Tree, Estimation of Distribution Algorithm

1 Introduction

The resource-constrained project scheduling problem (RCPSP) is an extension of the project scheduling problem (PSP), which has attracted an ever growing attention in recent years both from science and practice. Compared with conventional project scheduling problem, RCPSP is concerned with resource-constrained, each of which has a related duration and nonrenewable/renewable resource requirements.

Furthermore, in real-world problem, there are so many uncertainty and disruptions may be occurred, such as variation of duration time, which makes the problem much close to reality but more complex.

From[1], five kinds of approaches are summarized to deal with uncertainty in a scheduling environment where the evolution structure of the precedence network is deterministic: reactive scheduling, stochastic scheduling, scheduling under fuzziness, proactive/robust scheduling and sensitivity analysis. For reactive scheduling, it does not try to cope with uncertainty in creating the baseline schedule but revises or re-optimizes the baseline schedule when an unexpected event occurs. For proactive scheduling, the idea is to generate a robust baseline schedule by protecting it as well as possible against time and resource disruptions that may occur during project execution.

Proactive/reactive project scheduling that involves a proactive scheduling procedure for generating robust baseline schedule and the deployment of a reactive scheduling procedure that is to be invoked when the baseline schedule is disrupted and needs to be repaired. Commonly, it has tree stage process: (a) The generation of a precedence and resource feasible baseline schedule, (b) protecting the baseline schedule against disputations that may occur during project execution, and (c) deploying a reactive scheduling procedure to repair the baseline schedule during project execution when needed[2], [3].

In this paper, we propose an algorithm hybrid method of policy tree and EDAs for RCPSP with rework uncertainty, and if needed, the algorithm can be extended to solve other kinds of processing time uncertainty. We integrate proactive scheduling and reactive scheduling into one model, by using policy tree to model the stochastic problem, EDA is used to solve the problem by generate rule for policy tree with a linear function.

The paper is structured as follows. Section 2 makes the illustrations of EDAs and policy tree in previous literature. In section 3, a novel algorithm is proposed to generate rule for policy tree, which is used for generating proactive/reactive scheduling. Section 4 describes and formulates the problem of resource-constrained project scheduling with uncertainty as study case; experiments demonstrate the effectiveness of our algorithm by the comparison with other existing methods. Finally, section 5 presents the summary of work.

2 EDAs and Policy Tree

2.1 Estimation of Distribution Algorithm

Recently, there are growing interests in Estimation of Distribution Algorithms (EDAs) that build an explicit probabilistic model for the distribution of good solutions found so far and use the constructed model to guide further search behavior[4], [5].

In the previous literature, there are lots of researches on project scheduling under uncertainty. For deterministic RCPSP, exact algorithms such as mixed-integer linear programming or branch and bound can solve this problem. But for a complex problem or uncertainty project scheduling discussed here, heuristic methods or evolutionary algorithms are much more effective[6]. As we known, EDAs have been proved as very powerful evolutionary algorithm. Till now, EDA is one of the most effective methods to solve uncertainty project scheduling. In previous research, no matter for proactive scheduling or reactive scheduling, EDAs are mostly used for deciding the operation sequence or resource assignment priority. With the sampling methods, solution quality or scheduling stability can be calculated.

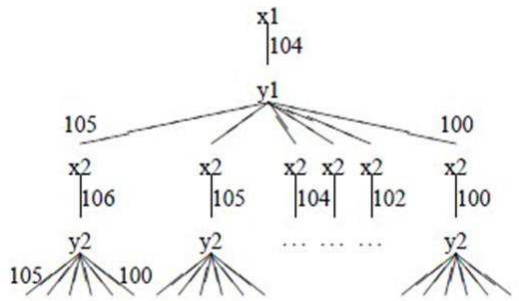


Figure 1. Example of policy tree in MSSCSP

But EDA still has its own demerit: when we trying to get one proactive scheduling and corresponding reactive scheduling for different kinds of disruptions, it is hard to decode in single EDA model. In next section, the model hybrid with policy tree will be introduced to overcome the shortage of EDA.

2.2 Policy tree

Policy tree was used in multi-stage stochastic constraint satisfaction problem (MSSCSP)[7]. In Figure 1, it shows the policy tree for example. A policy is a tree with nodes labeled with variables, starting with the first variable labeling the root, and ending with the last variable labeling the nodes directly above the leaves. Nodes labeled with decision variables have a single child, whilst nodes labeled with stochastic variables have one child for every possible value. Edges in the tree are labeled with values assigned to the variable labeling the node above.

The important one is that variables are set in the order. In a two stage stochastic CSP, the set of variable contains the first set of decision variables, the first set of stochastic variables, then the second set of decision variables, and finally the second set of stochastic variables.

Each path in the tree represents a different possible scenario and the values assigned to decision variables in that scenario. Each leaf node corresponds to a possible world and has an associated probability. The probability of one path is simply multiplied by each probability of the value given to the stochastic variables.

The policy tree can be viewed as one kind of decision tree, but in this paper, we will use this tree structure to model the scheduling problem with rework uncertainty, and generate the proactive scheduling and reactive scheduling dealing with disruption of rework.

3 Proposed Algorithm of PT-EDA

3.1 Overview of proposed PT-EDA

In Figure 2, the framework of our proposed PT-EDA is given. First step, one initial probability matrix is created, after that by sampling the probability matrix, we can get some individuals.

Then, we evaluate each individual by the average fitness of all schedules generated by the

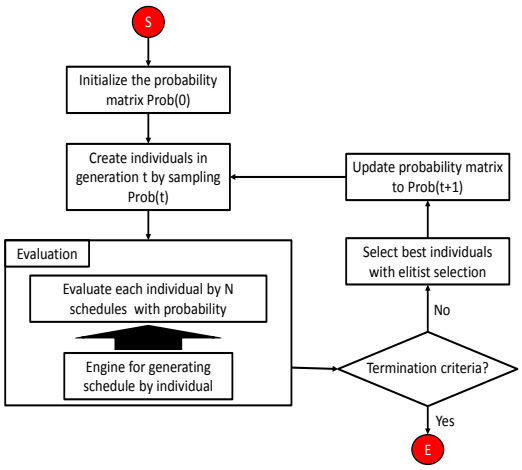


Figure 2. Framework of PT-EDA

individual with the scenario reduction. Next, the best individuals are selected based on fitness value. Subsequently, the selected individuals are used to update probability matrix. Finally, if the fitness of one individual reaches the termination criteria, it stops[8].

The key points of PT-EDA are: formulation of policy tree for proactive-reactive scheduling and method to generate policy tree. Two parts are discussed in next two subsections.

3.2 Rule for scheduling by policy tree

Based on policy tree shown in Figure 1, a new modeling of policy tree for reactive project scheduling is proposed. The illustration of policy tree for project scheduling problem with rework uncertainty is shown in Figure 3.

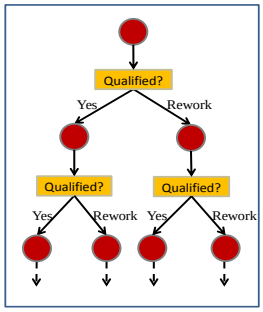


Figure 3. Policy tree

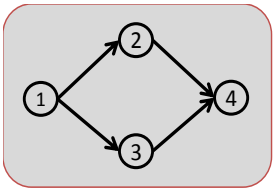


Figure 4. 4 activities

Similar with used in multi-stage, in Figure 3, we formulate policy tree for project scheduling with rework uncertainty as following:

(a) Take the multi-stage as "sequence", in each level of tree has one node (activity) and edge (rework or not).
(b) In each level, the value represents a pseudo-value for generating the activity id.
(c) After each decision value in the red circle, yellow rectangle followed represents the stochastic variable, which here means the corresponding activity's rework or not with probability.
(d) Each path from root to leaf represents one scenario with combined probability of all the activities.

For one path from root to leaf, it can be represented as one schedule under the corresponding scenario. The sequence also represents the resource assignment sequence (precedence relations also need to consider to make schedule). Firstly we assign the activity in the root node to the resource, and then we assign the activity in the node of second level, and so on. Take one simple example for explanation, in Figure 4, there are totally four activities (dummy activity omitted) named Activity 1, 2, 3 and 4. The resource requirements of four activities are 3 units, 1 unit, 1 unit and 3 units of resource R. The per-period-availability of resource R is 4 units. Suppose the sequence represented by one path from root to leaf is 1->3->2->4. Firstly we assign Activity 1, 3 units of R are holding and 1 unit of R remains, then we assign Activity 3, although only 1 unit of R is need, constraint of precedence relations cannot be violated and it can be operated only after activity 1 completed. With the same manner, take consider of constraint of precedence relations and resource capacity per-period, we can complete the schedule. This progress is shown in Figure 5.

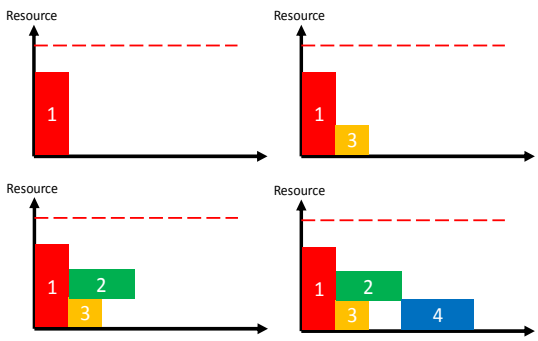


Figure 5. Generate schedule with one path

As shown in Figure 6, after one satisfy policy tree is given ("satisfy" here means the average makespan achieve a value or heuristic procedure finish), firstly we select the scenario (one path in the tree) with the highest probability as initial one and generating schedule as proactive scheduling. If activity j in this scenario needs rework, another branch of edge immediate after node j will be selected, and choose the highest probability path among all the paths left. We use this kind of path selection to deal with disruption and reactive scheduling.

3.3 Function to generate policy tree

We have known that RCPSP with rework uncertainty is NP hard problem[9]. Although we have model it by binary tree structure, it is still hard to solve it by B&B search. As the number of activities increase, the scenario will increase exponentially[10], which makes the problem impossible to get satisfy solution by ILOG solver with the acceptable calculation time.

There are two demerits: (1) For large problem, there are so many "paths", so that too many decision variables need to calculate (memory space problem); (2) With the number of decision variables increase, it becomes so many scenarios existing to check.

Not like the traditional heuristic methods (such as GA) by giving each node a gene, in our algorithm, we give each level a weight, which is viewed as one gene. Take the totally J activities for example, we can reduce length of chromosome from 2^{J+1}-1 down to 2J (J levels nodes and J levels edges), which will make the problem to be more practical to solve. As shown in Figure 7, it represents that we give each level a weight value, by which we calculate the value for each node.

As one example, as shown in Figure 8, if we want to get the value of C, similar with the

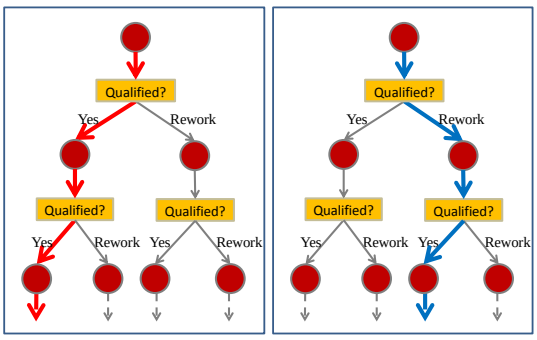


Figure 6. Path changeover after disruption

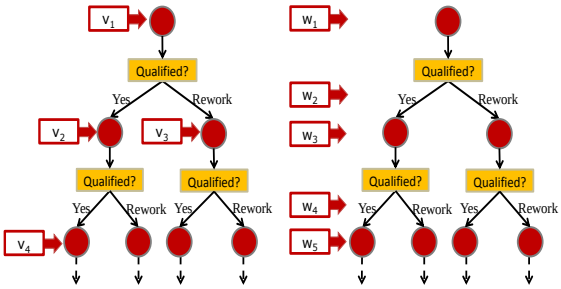


Figure 7. Two decode methods: value for node & weight for level of tree

Markov chain, it is determined by value of node A, node A qualified, node B and node B unqualified, in the other word, it is depend on all the decision variables and stochastic variables that preceding. In the project scheduling problem, from the view of resource assignment, when we try to select next activity to assign, the resource remaining and precedence relations must be taken into consider, which are influenced by the activities have been assigned before and whether they need rework. And here the weight of each level represents the degree of influence.

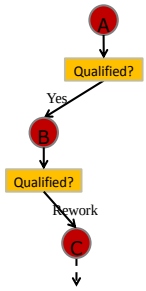


Figure 8. One path of policy tree

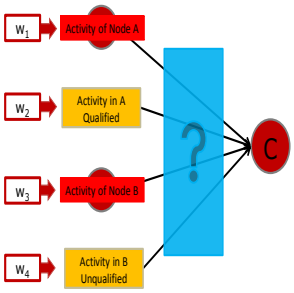


Figure 9. Determine by preceding variables

Like the theory of neural network (NN), if we can find the relationship among these variables, then we can get each value for nodes in the policy tree. The weights in chromosome can represent the relations. In Figure 9, it shows the relations among these variables.

First step, we generate integer number by involving all relevant stochastic variables and decision variables proceeding in linear function.

$$\alpha_{\bar{\omega}}(D,S,x_q) = \omega_q + \sum_{i\in\sigma_q}(\omega_i \times d_i) + \sum_{j\in\tau_q}(\omega_j \times s_j) \qquad (1)$$

where σ_q and τ_q denotes the set of all nodes and edges preceding, ω_q, ω_i and ω_j are weight values in the gene, d_i and s_j are values of decision and stochastic nodes.

Second step, we try to calculate a number between 1 and J, which is called pseudo-value in

our algorithm.

$$v(x_q) = \text{Lower} + \lfloor \alpha_{\bar{\omega}}(D,S,x_q) \rfloor \ mod \ (\text{Upper} - \text{Lower} + 1)$$
$$= 1 + \lfloor \alpha_{\bar{\omega}}(D,S,x_q) \rfloor \ mod \ \text{J} \tag{2}$$

where "Upper" and "Lower" are interval boundary, in RCPSP, it represents from 1 to J, and here means J.

Third step, we choose one activity by the mapping method from the assignment set. The assignment set contains all the activities can be executed with considering precedence relations.

4 Experiments

It is extremely difficult to get real world problem data on RCPSP. Fortunately there has been a significant amount of researches conducted on project scheduling problem, and one popular benchmarks problem data existing called PSPLIB (http://webserver.wi.tum.de/psplib/), which can be used for the comparison of different methods. In our study, we use two testing problems J10 and J30, which represented as medium and large problem.

4.1 Study case: project scheduling with rework uncertainty

The resource-constrained project scheduling problem (RCPSP) is to study how to allocate renewable/nonrenewable resource and schedule activities to minimize the whole project makespan. During project execution, however, project activities are subject to considerable uncertainty that may lead to numerous schedule disruptions, which comes from: activities may take more or less time than originally estimated, resources may become unavailable, ready times and due dates may have to be changed, etc.

Here we take consider of rework. It means that there is always one inspection after each activity completed. When the result of inspection shows the activity is unqualified, so that this one activity needs an extra time to reproduce. When the inspection is qualified, the next activity in schedule can be operated. In the other word, it's one special case of time variation with two conditions. For each activity, there are two states including qualified and unqualified, so that totally have 2^J different kinds of probable states combination when the operation sequence fixed. One probable combination is called as one scenario[11], for example, all activities are qualified is one ideal scenario.

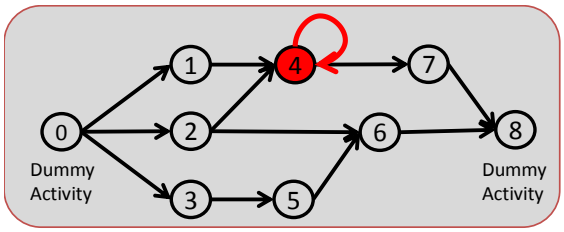


Figure 10. RCPSP with rework

We take this application for our research with the following reasons: Firstly, it's very practical problem in real world, especially for the problem mainly operated by labor. Not like machines, due to the different level skills of workers, the product flaw is frequently occurred. Secondly, this problem can be viewed as one kind of time variation. We can extend this problem to other different kinds and can be easily solved with the same manner by our proposed algorithm.

Figure 10 illustrates a simple example with 9 activities (including two dummy activities j=0 and j=J+1). Compared with traditional RCPSP, the red circle of activity 4 represents the probability of rework. Four kind of renewable resource are considered, and resource requirements, durations and corresponding rework probability are listed in Table 1. Table 2 shows the precedence relationship of each activity by its successors.

Table 1. Resource and rework probability

ID	Dura	Resource				Prob.
		R1	R2	R3	R4	
1	0	0	0	0	0	0
2	8	4	0	0	0	20%
3	4	10	0	0	0	10%
…	……	…	…	…	…	……
Capacity		12	13	4	12	

Dura=Duration; Prob.=Probability

Table 2. Precedence relationship

ID	#successors	Successors			
1	3	2	3	4	
2	1	5			
3	2	5	7		
…	……	…	…	…	…

#successors=no. of successors

Generally, the RCPSP with rework uncertainty

can be stated as follows[12], [13]. A project is composed of J activities. The set of renewable resource is referred as K^r, and the per-period-availability is assumed to be constant R_k^r. Precedence relationship for each activity j=1,2,…,J is defined by sets of its immediate predecessors P_j, which indicate that the activity j cannot be executed before any of its predecessors $i \in P_j$ is not completed yet. For activity j=1, 2, …, J, the duration is d_j. In this duration, r_j^k units of renewable resource $k \in K^r$ is consumed in each period of the non-preemptable duration. Each activity j=1, 2, …, J has a corresponding rework probability of $Prob_j$. The dummy activities j=0 and j=J+1 represent the start and end of the project, respectively. The dummy activities do not request any resource and have zero duration. The set of all activities can be denoted as $J^+ = \{0, \ldots, J+1\}$. The objective is to minimize average makespan of different scenarios with probability.

The RCPSP can be formulated as follows:

Objective:

$$Minimize \sum_{\theta=0}^{2^J-1} \left(\prod_{j=1}^{J} [\mathrm{Prob}_j \times (1 - X_{j\theta}) + (1 - \mathrm{Prob}_j) \times X_{j\theta}] \times \mathrm{F}_{J+1,\theta}\right) \quad (3)$$

Subject to:

$$F_{i,\theta} \le F_{j,\theta} - d_{j,\theta},\ j = 1, 2, \ldots, J+1, \mathrm{i} \in \mathrm{P}_j \quad (4)$$

$$d_{j,\theta} = (1 + X_{j,\theta}) \times d_j \quad (5)$$

$$F_{j,\theta} \ge 0,\ j = 1, 2, \ldots, J+1 \quad (6)$$

$$\sum_{j \in A_\theta(t)} r_{j,k} \le R_k, k \in K, t \ge 0 \quad (7)$$

$$A_\theta(t) = \{\mathrm{j} \in \mathrm{V} \mid \mathrm{F}_{j,\theta} - d_{j,\theta} \le t < F_{j,\theta}\} \quad (8)$$

$$X_{j,\theta} = \begin{cases} 0 & if\ \ \theta \in (0, 2^{J-1} + 1 - \mathrm{j}) \cup (2^J - j, 2^J + 1) \\ 1 & else \end{cases} \quad (9)$$

$$1 \le \theta \le 2^J \quad (10)$$

where $F_{i,\theta}$ and represents completion time of activity i in scenario θ, and $d_{i,\theta}$ represents processing time of activity i in scenario θ. $A_\theta(t)$ means the assignment set at time t. $X_{i,\theta}$ equals 1 if activity i in scenario θ needs rework, else equals 0.

Table 3. Comparison on scheduling makespan

P	Scheduling	GA	EDA	PT-EDA
J10	Baseline	18.2	18.1	18.5
	Reactive	22.3	21.9	21.3
J30	Baseline	53.6	53.2	54.1
	Reactive	60.7	60.3	58.8

Equation (3) shows the objective of project scheduling, to minimize the weighted makespan of each scenario. Equation (4) (5) and (6) represent the constraints of precedence relations, an activity can be performed must after all its predecessors finished, where $F_{j,\theta}$, $d_{j\theta}$ represent the finish time and duration time of activity j in the scenario θ. Equation (7) and (8) represent the constraints of capacity of each resource, where $A_\theta(t)$ indicates the set of activities which is operating at time t in the scenario θ. Equation (9) and (10) represent whether activity j needs rework from the scenario θ=1 to 2^J.

4.2 Experiment results

In order to demonstrate the performance of our propose algorithm PT-EDA. Two experiments are performed: First, make the comparisons with deterministic RCPSP on the objective value of makespan and stability, which is to prove that, our proactive scheduling can get almost the same result with deterministic ones, but overtake on the reactive scheduling, which means our algorithm can produce good schedules with better tolerance with uncertainty in execution stage[14].

In order to equally make comparisons, GA and our proposal each runs 30 times with the same 1000 generations. Compared with GA, our proposal gets better solution with less value of the best solution, and better calculation stability with less mean value.

In Table 3, it demonstrates the efficiency of reactive scheduling. From the results, if we do not consider the reactive scheduling for dealing with disruptions, we can get less makespan of baseline scheduling. However, based on this baseline scheduling, reactive scheduling will get larger makespan. By our proposal, although baseline is worse, it is so close and gets better result on reactive scheduling.

From Figure 11, we can see the proposal can achieve better results on the convergence than

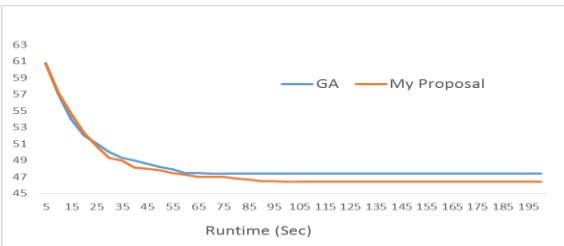


Figure 11. Comparison of Convergence

traditional GA-based algorithm. With the same calculation time, PT-EDA convergences much more quickly than GA and gets less makespan.

5 Conclusion and Future Work

In this paper, we have proposed an algorithm called PT-EDA to solve resource-constrained project scheduling problem (RCPSP) with considering rework with probability. In PT-EDA, we use policy tree to model the stochastic project scheduling problem, and in order to avoid the memory problem associated with large tree, a linear function is used to generate policy tree, which is acted as the rule to generate dynamic scheduling for solution. Experiments results based on benchmarks (PSPLIB) and comparisons with some existing heuristic algorithms to demonstrate our approach with high effectiveness and tolerance with uncertainty in both planning and execution stage.

From the previous sections, we have known that PT-EDA is one general method that can be applied to other experiments, such as FJSP or any scheduling problems. Because of the time limitations, we still need to make more experiments on other problems to demonstrate our method can achieve better results on different kinds of reactive and proactive scheduling problems.

In the future work, multi-mode can be added because for each activity has several execution modes, each of which has a related duration and nonrenewable/renewable resource requirements. This scheme makes the problem much more close to reality and more meaningful to make the schedule. Different mode leads to different rework probability: often high resource requirement will lead to lower rework probability, we should make tradeoff between makespan and resource usage.

References

[1] W. Herroelen and R. Leus, "Project scheduling under uncertainty: Survey and research potentials," *Eur. J. Oper. Res.*, vol. 165, no. 2, pp. 289–306, 2005.

[2] E. L. Demeulemeester and W. Herroelen, *Robust project scheduling*, vol. 3. Now Publishers Inc, 2011.

[3] W. Herroelen and R. Leus, "Robust and reactive project scheduling: a review and classification of procedures," *Int. J. Prod. Res.*, vol. 42, no. 8, pp. 1599–1620, 2004.

[4] P. Larranaga and J. A. Lozano, *Estimation of distribution algorithms: A new tool for evolutionary computation*, vol. 2. Springer, 2002.

[5] M. Pelikan, D. E. Goldberg, and F. G. Lobo, "A survey of optimization by building and using probabilistic models," *Comput. Optim. Appl.*, vol. 21, no. 1, pp. 5–20, 2002.

[6] L. Wang and C. Fang, "An effective estimation of distribution algorithm for the multi-mode resource-constrained project scheduling problem," *Comput. Oper. Res.*, vol. 39, no. 2, pp. 449–460, 2012.

[7] T. Walsh, "Stochastic constraint programming," in *ECAI*, 2002, vol. 2, pp. 111–115.

[8] X. Hao, X. Chen, H. W. Lin, and T. Murata, "Cooperative Bayesian Optimization Algorithm: A Novel Approach to Simultaneous Multiple Resources Scheduling Problem," in *Innovations in Bio-inspired Computing and Applications (IBICA), 2011 Second International Conference on*, 2011, pp. 212–217.

[9] S. Hartmann and D. Briskorn, "A survey of variants and extensions of the resource-constrained project scheduling problem," *Eur. J. Oper. Res.*, vol. 207, no. 1, pp. 1–14, 2010.

[10] S. Prestwich, S. A. Tarim, R. Rossi, and B. Hnich, "Evolving parameterised policies for stochastic constraint programming," in *Principles and Practice of Constraint Programming-CP 2009*, Springer, 2009, pp. 684–691.

[11] S. Manandhar, A. Tarim, and T. Walsh, "Scenario-based stochastic constraint programming," *Arxiv Prepr. Arxiv09053763*, 2009.

[12] P. Brucker, A. Drexl, R. Möhring, K. Neumann, and E. Pesch, "Resource-constrained project scheduling: Notation, classification, models, and methods," *Eur. J. Oper. Res.*, vol. 112, no. 1, pp. 3–41, 1999.

[13] R. Kolisch, "Serial and parallel resource-constrained project scheduling methods revisited: Theory and computation," *Eur. J. Oper. Res.*, vol. 90, no. 2, pp. 320–333, 1996.

[14] S. Hartmann, "A competitive genetic algorithm for resource-constrained project scheduling," *Nav. Res. Logist. Nrl*, vol. 45, no. 7, pp. 733–750, 1998.

Human Detection from Surveillance Cameras with Considering the Influence of Shadows

Ryo Yonekawa† **Masao Izumi†**

†Graduate School of Engineering, Osaka Prefecture University

Abstract

Nowadays, there are huge numbers of surveillance cameras installed in the world everywhere. And many researches cover surveillance cameras installed outdoors. Surveillance cameras installed outdoors receive considerable environmental influence. Particularly, shadows raise the incidence of occlusions of the object regions. In this paper, we propose the novel method of human detection with considering the influence of shadows. At first, our proposal method performs background subtraction using normalized distance. Next, it updates background using mixed Gaussian distribution dynamically. Afterwards, it changes the size of the extracted object regions into four phases and performs human detection by using HOG (Histograms of Oriented Gradients) features to classify detection windows into human or non-human category with SVM (Support Vector Machine). Experimental results show that the proposed method is effective to extract correct object regions with surveillance cameras installed outdoors.

Keywords: background subtraction, normalized distance, mixed Gaussian distribution, human detection, Histograms of Oriented Gradients, SVM

1 Introduction

Surveillance cameras have been installed in the world everywhere. They have been installed in such areas as public accommodations, stations, buildings, public loads, etc, where a large number of people come into or go out. Main purpose of using surveillance cameras is crime prevention. However, it is difficult to keep watching the surveillance camera images with human eyes because the abnormal situation rarely happens. Thus we hope the system can automatically detect abnormalities and inform a watchman. To that end, a new image processing technology to detect target objects only from images is necessary. In this study, we perform the experiment for the purpose of detecting only humans precisely from images of fixed cameras installed outdoors.

2 Outline of the proposed method

This study covers surveillance cameras installed outdoors. Therefore, the camera is affected by the dim change of the background and sunshine condition. Because the shadow in particular makes it easy to generate the occlusion of the object regions, it is an obstacle in securing correct object regions. Many conventional methods can adapt to the dim change of the background, but were not able to remove the shadow completely. Accordingly, our proposal method performs background subtraction using normalized distance [1] that is effective for the shadow removal. Next, it updates the background with mixed Gaussian distribution [2][3] to be adapted to the dim change of the background. And it removes the noise of the extracted object regions, performs labeling in each domain and performs human detection using HOG features to classify detection windows into human or non-human category with SVM. We show a flow of the proposed method in figure 1.

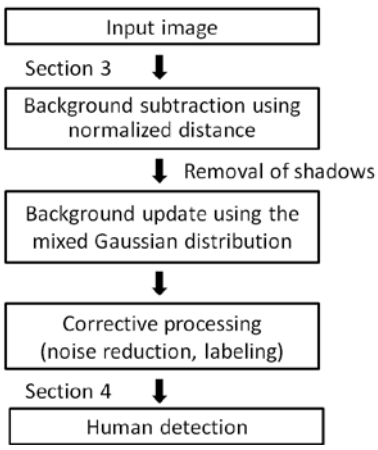


Figure 1. Flow of the proposed method

3 Object extraction in consideration of the influence of the shadows

In this section, we suggest the method that combines mixed Gaussian distribution with background subtraction using normalized distance.

3.1 Background subtraction using normalized distance

As the background subtraction that can secure the object regions except the shadows, there is a method using normalized distance [1]. In local area, this method calculates normalized distance from a background image and an input image. When each luminosity value of the background image and that of the input image are correlative to each other, the normalized distance is close to 0. The normalized distance D_N is calculated in the next formula.

$$D_N(i,b) = \left| \frac{i}{|i|} - \frac{b}{|b|} \right| \quad - ①$$

Where, "i" is a vector of local area's luminosity values of the input image. "b" is a vector of local area's luminosity values of the background image. $|i|, |b|$ is calculated with the square root of sum of squares of local area's luminosity value respectively. Shadows tend to darken the shadowed area with the certain range of proportions. In other words, the ratio of luminosity between shadowed and non-shadowed in each local area keeps constant. For that reason, lower value of normalized distance tends to be shadowed area, and higher value of normalized distance tends to be non-shadowed area. (Figure 2.)

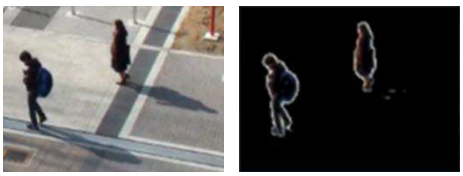


Figure 2. Background subtraction using normalized distance

3.2 Background update using the mixed Gaussian distribution

Background subtraction using normalized distance needs to fix the background image which we acquired from camera images beforehand. For that reason, this method is very weak to camera shake and the dim change of the background. Thus we suggest the method that combines mixed Gaussian distribution [2][3] with background subtraction using normalized distance. (Figure 3.) The mixed Gaussian distribution is used to update dynamically the background image.

The following is a brief flow of the handling of background update using the mixed Gaussian distribution. Refer to documents [2] for the details of the method.

Step1
This method makes mixed Gaussian distribution from a pixel value of the constant section. And, it divides the Gaussian distribution into plural Gaussian distribution by EM algorithm. It calculates a collation degree of each Gaussian distribution on this occasion.

Step2
If a distance between a pixel value of the input and the mean of Gaussian distribution is less than 2.5 times of standard deviation of Gaussian distribution, it considers the pixel matches the Gaussian distribution.

Step3
If a distance between a pixel value of the input and the mean of Gaussian distribution is less than 2.5 times of standard deviation of Gaussian distribution, it considers the pixel matches the Gaussian distribution. If Gaussian distribution is found, it estimates the background of the pixel than the collation degree of the Gaussian distribution.

Step4
It updates the collation degree of each divided Gaussian distribution.

(a)input image (b)normalized distance (c)proposal method

Figure 3. Proposal method

3.3 Corrective processing (noise reduction, labeling)

After the processing mentioned above, the left noises can be eliminated using expansion reduction. Besides, we perform labeling in order to

separate extracted regions into correct object regions.

4 Human Detection

HOG expresses the distribution of the intensity gradient in local area, it can accommodate the change of the person's shape. SVM is the learning model that is excellent in recognition performance and is one of the pattern identification methods.

4.1 Summary of human detection using HOG and SVM

The human detection using HOG [4] is divided into a learning process and a detection process. In the learning process, we calculate HOG from a large quantity of samples and SVM learns a shape characteristic of the target for the detection. In the detection process, a detection window to detect HOG performs raster scan on an image. The detection window changes size and performs multiple raster scan. SVM judges whether HOG which is calculated by a detection window is human or not.

4.2 Calculation of HOG

HOG[4] is calculated by the following flows. In addition, we show the relations among an input image, blocks, cells and pixels be used for calculation of HOG in following figure 4.

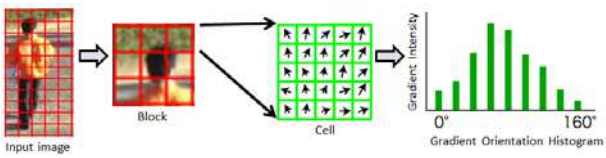


Figure 4. Block and cell to use for calculation of HOG

Step1 Calculation of gradient intensity and gradient orientation

(x, y) is the coordinate of images. Luminosity gradient f_x, f_y is calculated by the next formula.

$$\begin{cases} f_x(x,y) = L(x+1,y) - L(x-1,y) \\ f_y(x,y) = L(x,y+1) - L(x,y-1) \end{cases} \quad - ②$$

In addition, gradient intensity and gradient orientation are calculated by the next formula.

$$\begin{aligned} &\text{Gradient Intensity} : m(x,y) = \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \\ &\text{Gradient Orientation} : \theta(x,y) = \tan^{-1}\frac{f_y(x,y)}{f_x(x,y)} \end{aligned} \quad - ③$$

Step2 Making of the gradient orientation histogram of the luminosity

The calculated gradient orientation becomes $0^\circ \sim 360^\circ$. However, because it is not necessary to consider the direction of the gradient orientation, we convert it into $0^\circ \sim 180^\circ$ and divide it into 9 directions by 20° steps. Moreover, add gradient intensity of the luminosity in each calculated directions and make gradient orientation histogram.

Step3 Normalize each block

We normalize each block by using the following formula.

$$v(n) = \frac{v(n)}{\sqrt{\left(\sum_{k=1}^{q \times q \times N} v(k)^2\right) + \epsilon}} \qquad (\epsilon = 1) \qquad N: \text{Gradient Orientation Number} \quad - ④$$

ε is a coefficient to prevent a denominator from becoming 0.

5 Experiment and consideration

We show an inspection result based on the procedure of Section 3 and Section 4.

5.1 Experimental environment

There were three camera images of each 50 seconds which varied in the photography date, a sunshine condition, and a congestion condition. Besides, we used HOG data which SVM learned for human detection. This data was prepared for by OpenCV2.3.1 [5].

The followings show the condition of this experiment.

- photography scale : 30 meters - 35 meters in width
- camera images① : influence of shadows and trees, camera images② : influence of the shadow, camera images③ : no influence of shadows and trees
- scene number : 3 (50 second))
- resolution : 640x480
- frame rate : 6 frame/second
- threshold 2.7 of normalized distance
- resized size of object regions for human detection:80x140,90x140,100x140,110x140
- HOG calculation parameter : (cell size 8x8 , block size 16x16 , block movement size 8x8 , the number of edge directions 9)

5.2 Experimental results

As a preliminary experiment, we test detection precision of HOG data which was prepared for by OpenCV2.3.1 [5]. We found the rate of detection for 1,033 pieces of positive samples (Figure 6) in the preliminary experiment. Following table 1 shows the result.

Table1 Preliminary experiment

Preliminary experiment	Positive sample	Rate of detection	Detection error
	1033	91.7%	8.3%

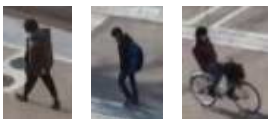


Figure 6. Example of positive sample

Next, following table 2 shows the result which compared among the mixed Gaussian distribution, normalized distance and proposed method.

Table2 Experimental results

		Mixed Gaussian	Normalized distance	Proposal method
①	Number of region	3356	2098	1442
	Rate of human detection	36.5%	55.7%	70.1%
②	Number of region	2637	1112	1102
	Rate of human detection	37.5%	81.7%	76.8%
③	Number of region	2551	1984	1828
	Rate of human detection	72.3%	80.2%	70.9%

5.3 Consideration

In case of camera images①, the object regions extracted by mixed Gaussian distribution have the low rate of human detection. It is because an object domain received the considerable influence of shadows. Further, background subtraction using normalized distance is necessary to fix a background image. Hence, shaking of trees caused the false detection. On the other hand, the proposal method was able to get the highest rate of human detection because it could cope with the influence of shadows and trees. In the case of camera images② and camera images③, the proposal method gets high rate of human detection. The proposal method is effective in surveillance cameras installed outdoors which cannot avoid environmental changes.

6 Conclusion

We performed background subtraction using normalized distance that is effective for the shadow removal in this study. This method reduced the occlusion of object regions which shadows caused and made it possible to secure a more correct object domain. Surveillance cameras installed outdoors is affected by the dim change of the background and sunshine condition. Therefore, we have proposed the method that combined mixed Gaussian distribution (which was one of methods to update a background dynamically) with background subtraction using normalized distance. As a result of this experiment, we found that the proposal method is effective in surveillance cameras installed outdoors which cannot avoid environmental changes.

References

[1] Nagaya, S.,Miyatake, T.,Fujita, T,.Ito, W., Ueda, H.,"Moving Object Detection by Time-Correlation-Based Background Judgement Method", IEICE, vol.J76-D- Ⅱ, no.8, pp-1562-1571, 1993 (in Japanese).

[2] N. Friedman, S.Russell, "Image segmentation in video sequences : A probabilistic approach", 13th Conference on Uncertainly in artificial intelligence, 1997.

[3] Shimada, A., Arita, D., Taniguchi, R.," Increment and Decrement of Gaussians in Adaptive Mixture-of-Gaussian Background Models", IEICE, MIRU2006, pp-746-751, 2006-07 (in Japanese).

[4] N.Dalal, B.Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE CVPR, pp. 886-893, 2005. on Computer Vision and Pattern Recognition (CVPR' 01) - Volume 1, 2001.

[5] OpenCV, URL:http://opencv.willowgarage.com/wiki/ , Accessed on March.7, 2013.

Automatic Update Method of WiFi Indoor Positioning Model using Pedestrian Dead-Reckoning Technique

Taniuchi Daisuke[†] **Maekawa Takuya**[†]
[†]Graduate School of Information Science and Technology, Osaka University
{taniuchi.daisuke,maekawa}@ist.osaka-u.ac.jp

Abstract

WiFi technology is very popular and has been implemented in public areas such as hospitals and train stations. So, many researchers have attempted to construct indoor positioning systems by utilizing WiFi access points (APs) with small deployment costs. WiFi based indoor positioning systems rely on a public infrastructure that is not controlled by the user or the developer of the systems. So, when one WiFi AP is removed or moved, or layout changes, recalibration of the positioning systems is required. To reduce the cost, this paper proposes a new method for maintaining the positioning system. In our proposed method, we firstly track the user, who spends a lot of time in a given environment, with pedestrian dead reckoning techniques. At the same time we obtain WiFi scan data from a mobile device possessed by the user. With the scan data and the estimated coordinates, we can automatically create a pair consisting of a scan and its corresponding indoor coordinates during the user's daily life. Then, we periodically update signal strength fingerprints by using the information to cope with the instability of WiFi based positioning methods caused by changing environmental dynamics.

Keywords: Indoor positioning, fingerprinting, pedestrian dead reckoning

1 Introduction

Due to the recent proliferation of cheap and small sensors, many researchers in the wearable computing research field have employed body-worn sensors such as accelerometers and orientation sensors to track the movement trajectory of a person indoors by detecting steps, and estimating stride lengths and the directions of motion [1]. This methodology is called Pedestrian Dead Reckoning (PDR).

On the other hand, WiFi technology is very popular and has been implemented in public areas such as hospitals and train stations. So, many researchers have attempted to construct indoor positioning systems by utilizing WiFi access points (APs) with small deployment costs. Fingerprinting techniques based on WiFi have usually been employed to measure indoor positions [2]. However, WiFi based indoor positioning systems rely on a public infrastructure that is not controlled by the user or the developer of the systems. So, when one WiFi AP is removed, for example, recalibration of the positioning systems is required. Also, in a dynamic environment caused by layout changes, the WiFi signal strengths measured in the positioning phase may deviate significantly from those stored in the database.

To cope with the problem, we should manually re-collect fingerprints. In this paper, we propose a new approach that automatically and periodically re-collects fingerprints and uploads them to the fingerprint database. We assume that, in an indoor environment such as a hospital or office, special users, such as workers who spend a lot of time in the environment, wear various kinds of sensors such as accelerometers and orientation sensors, and we attempt to track these workers with the PDR technique. We also assume that mobile devices carried by the workers continually scan WiFi signals. So, by combining the WiFi scan data and the estimated coordinates obtained during 'walk' activities in their daily lives, we can automatically create a pair consisting of a WiFi scan and corresponding indoor coordinates while placing only a small burden on the workers. The WiFi fingerprints are automatically updated by using the information during the workers' daily routine. With this approach, ordinary users (customers and other workers) who do not wear rich sensors can benefit from the continually updated WiFi fingerprints.

2 Proposed Method

2.1 Overview

Fig. 1 provides an overview of our approach. Here we assume that few Bluetooth beacons are

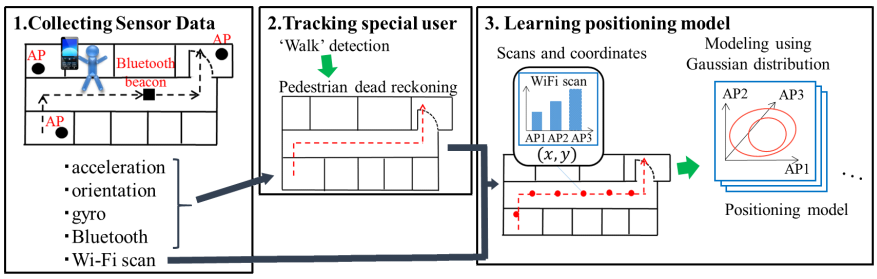


Figure 1. Overview of our proposed approach.

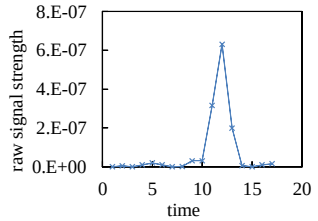


Figure 2. Example Bluetooth signal strength time series (2 Hz sampling).

installed in an indoor environment as shown in the left portion of Figure 1. The input of our method consists of sensor data obtained from mobile devices of special users (e.g., workers in a particular environment). Our method has two main components: *Tracking special user* and *Learning positioning model*. We explain them briefly below.

[Tracking special user]

We track the special user by using sensor data from the user. We achieve this by integrating techniques proposed in existing wearable and mobile computing studies. With the computed trajectory and the WiFi scan data obtained with a mobile device carried by the user, we can make pairs consisting of WiFi scan data and corresponding indoor coordinates. The created pairs are stored in the WiFi scan database. So, the database is continuously updated during the user's daily life.

[Learning positioning model]

We periodically update the WiFi positioning model by using the data stored in the database. We learn the features of the received signal strengths at each grid cell in the map of the environment. So, we aggregate scans within each grid cell. For each grid cell, we model the received signal strengths from the APs by using the scans in the cell.

We describe the procedure in detail below assuming that a device with sensors is attached to the user's body. Note that, in the following, we explain how we compute the user's trajectory of that after the user passed by the beacon. We can compute the trajectory before the passing or between the passings (if there are two or more beacons in the environment) in almost the same manner.

2.2 'Walk' detection

We first detect 'walk' activities (i.e., detect start and end times of walking) in the user's daily life by using acceleration and gyro sensor data obtained from the user (about 16 Hz). We employ the existing activity recognition methods proposed in wearable and pervasive computing studies [3; 4] that are based on supervised machine learning techniques.

2.3 Obtaining absolute coordinates

With the above procedure, we can find the start and end points (times) of a 'walk.' With the Bluetooth scan data obtained during the 'walk,' we can determine when the user passed by the Bluetooth beacon. When the user passes by the beacon, the Bluetooth signal strength gradually increases and then decreases as shown by the data in Fig. 2, which were obtained when an experimental participant passed by a beacon. So, when the value exceeds a certain threshold and becomes a local maximal value, we determine that the user passed by the beacon. Note that, when we cannot find Bluetooth signal data that exceed the threshold during walking, we discard the walking sensor data and do not use them in the following procedure.

2.4 Tracking with PDR

To track the user's trajectory, we employ a particle filter [5] that is usually used to estimate the states of non-linear systems. We explain the procedures required for the tracking.

[Initialization]

We first set particles at the beacon's coordinates. Since there may be small errors related to Bluetooth beacon based positioning, we randomly scatter several particles around the beacon's coordinates generated from a bivariate Gaussian distribution. We then determine the directions in which the particles move by using orientation sensor data. Here we assume that a beacon is installed at a place where the user's direction of movement is restricted (e.g., corridor). So, we preset the probable directions when a user passes the beacon in advance. Then, we determine in which direction the user is heading by using orientation data. We adopt this approach because the orientation sensor data obtained indoors in-

clude very large errors due to the presence of electric appliances and metals.

[Determining stride and direction]

We determine the speed and direction of the user's movement. First, we detect steps by using acceleration sensor data. We combine the acceleration data for three axes and simply detect steps by counting the number of times the gravity acceleration value is crossed by the combined signals. For each step, we assume that the user (particle) moves the length of a stride. In the sampling process, we generate several particles from each particle by randomly changing their length of stride by using a Gaussian distribution with a predefined mean because the length of stride may change for each step. At the same time, we also determine the direction of the user (particle) by using gyro sensor data because orientation sensor data are unreliable indoors. Since we know the direction of movement when the user passed the beacon, we can compute the direction after the time of passing by using relative angular speed data (gyro sensor data). We also randomly change the direction for each particle by using a Gaussian distribution to cope with its sensor data errors.

[Clustering]

As above we move particles from the time at which the user passed the beacon to the point (time) when the walking ends. When a particle collides with obstacles in the environment, the particle is discarded. If two or more particles survive at the time at which the walking ends, we should select one of them and employ the trajectory of the selected particle as the user's trajectory. We cluster the coordinates of the surviving particles at the end time with the X-means algorithm [6], and find the largest cluster. Then we select the particle closest to the centroid of the cluster. Also, when only one particle survives at the end of the walking, we simply determine the trajectory of the particle as the user's trajectory. When no particle survives at the end time, we discard the walking and do not use it in the following procedure.

2.5 Learning positioning model

Firstly, we aggregate scans within each predefined grid cell in the environment and learn the parameters of the signal strength distribution in the cell for each AP. Previous studies support the view that the WiFi received signal strength obeys a Gaussian distribution [7; 8] when the sample size is sufficient, and so we also use a Gaussian distribution to model the signal strength from each AP at each cell and its probability density function is represented as follows.

$$f(x_i, \mu_{i,n}, \sigma^2_{i,n}) = \frac{1}{\sqrt{2\pi\sigma^2_{i,n}}} \exp\left(-\frac{(x_i-\mu_{i,n})^2}{2\sigma^2_{i,n}}\right),$$

where x_i is the observed signal strength from the ith AP, and $\mu_{i,n}$ and $\sigma^2_{i,n}$ are the mean and variance of a Gaussian corresponding to the ith AP at the nth grid cell, respectively. So, we compute $\mu_{i,n}$ and $\sigma^2_{i,n}$ from the scans in the cell.

As described above, we can learn the signal strength distribution from the ith AP at the nth cell. So, we can compute the likelihood with which a WiFi scan x is observed at the nth cell as follows.

$$p(x, \lambda_n) = \sum_{i \in x} f(x_i, \mu_{i,n}, \sigma^2_{i,n}),$$

where the parameters of Gaussians related to the nth cell are collectively represented by λ_n.

In the positioning phase, we compute the likelihood of each cell (λ_n) for an observation x obtained at unknown coordinates. As with the kNN approach, we obtain cells with the top-k likelihood values and compute the weighted average of the cell coordinates. (The weight corresponds to the likelihood value.) The average coordinates become the estimated coordinates. ($k = 3$)

3 Evaluation

3.1 Data set

The sensor data were obtained on one floor of our graduate school building as shown in Fig. 3. ($1m \times 1m$ cell size in our implementation) On the first day, we collected WiFi fingerprints at the training points shown in Fig. 3. Fig. 3 also shows the position of a Bluetooth beacon. We then iterated the following procedures every day for 28 days: (1) An experimental participant walked around the floor at various times with a Google Galaxy Nexus smartphone attached to his waist. He randomly selected start and end coordinates on the floor and walked according to those coordinates. He repeated the walk twenty times a day. Then, pairs consisting of a WiFi scan and its indoor coordinates were uploaded to the WiFi scan database. (2) We reconstructed the indoor positioning model with the data stored in the database. (3) We collected WiFi scans at the test points shown in Fig. 3 and computed the positioning accuracy for the test data.

3.2 Evaluation methodology

We tested the kNN method ($k = 3$) in addition to our method. To investigate the effectiveness

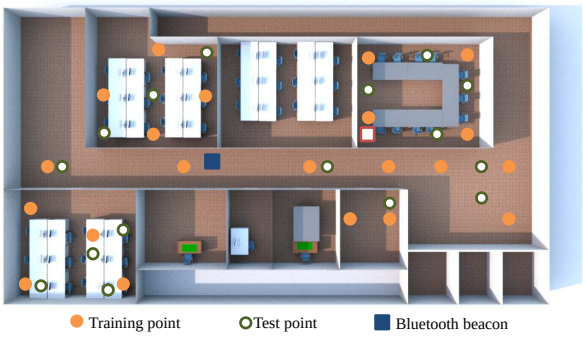


Figure 3. Floor plan of experimental environment (29.8m×16.3m).

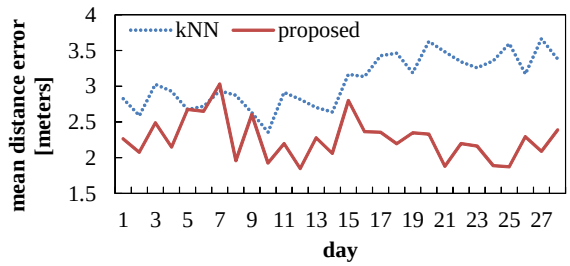


Figure 4. Transitions of accuracies related to kNN and our methods.

of our method, we virtually reduce the signal strengths from APs located outside of the floor after the 15th day on the assumption of some conversion of outer walls. We can observe stable signals in the environment from five APs located outside of the floor. We randomly reduce the signals using a Gaussian distribution with a 15 dBm mean based on our observation and existing studies [9; 10].

3.3 Results

Fig. 4 shows the transitions of accuracies related to the kNN and our methods when we reduced the signal strengths from the APs located outside the floor on and after the 15th day. Although the positioning performance of our method was not stable in the early stage, it provided good accuracy after sufficient numbers of scans were uploaded to the scan database. It greatly outperformed the kNN method. Our method created fingerprints densely placed in the environment by using the walking of the participants, and the dense fingerprints achieved good accuracies.

The mean error distance of the kNN method increased greatly (about 0.6 meters) after the reduction on the 15th day. Although the number of APs located outside the floor (five) was much smaller than that located on the floor (seventeen), the effect was significant. Immediately after the reduction, the mean distance errors with the proposed method also increased. However, the mean distance error gradually decreased because the new scans were continually uploaded by the participant.

4 Conclusion

In this paper we proposed a new approach for maintaining a WiFi indoor positioning model with respect to changing environmental dynamics with the help of body-worn sensors attached to specific users. We attempted to track the user with PDR techniques, and at the same time we obtained WiFi scan data from the user. With the scan data and the estimated coordinates, we automatically created a pair consisting of a scan and its corresponding coordinates, and updated the positioning model by using the information.

In our experimental evaluation, we confirmed that our method could cope with the sudden changes in signal strengths from APs.

References

[1] C. Randell, C. Djiallis, and H. Muller. Personal position measurement using dead reckoning. In *ISWC 2005*, pages 166–173, 2005.

[2] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, et al. Place lab: Device positioning using radio beacons in the wild. In *Pervasive 2005*, pages 116–133, 2005.

[3] L. Bao and S.S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive 2004*, pages 1–17, 2004.

[4] T. Maekawa and S. Watanabe. Unsupervised activity recognition with user's physical characteristics data. In *Int'l Symp. on Wearable Computers*, pages 89–96, 2011.

[5] A. Doucet, N. De Freitas, and N. Gordon. *Sequential Monte Carlo methods in practice*. Springer Verlag, 2001.

[6] D. Pelleg, A. Moore, et al. X-means: Extending k-means with efficient estimation of the number of clusters. In *ICML 2000*, volume 1, pages 727–734, 2000.

[7] W.-P. Chen, J. C. Hou, and L. Sha. Dynamic clustering for acoustic target tracking in wireless sensor networks. *IEEE Transactions on Mobile Computing*, 3(3):258–271, 2004.

[8] E. Chan, G. Baciu, and S. C. Mak. Using Wi-Fi signal strength to localize in wireless sensor networks. In *International Conference on Communications and Mobile Computing*, volume 1, pages 538–542, 2009.

[9] A. P. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri. Real-world environment models for mobile network evaluation. *IEEE Journal on Selected Areas in Communications*, 23(3):622–632, 2005.

[10] J. C. Stein. Indoor radio WLAN performance part II: Range performance in a dense office environment. Intersil Corporation, 1998.

Indoor Positioning System using Ultrasonic Sensors as a Position Information Infrastructure for a Wide Area

Chokatsu Yara† **Masaki Akiyama†** **Shigenori Ioroi†‡** **Hiroshi Tanaka†‡**

†Graduate School of Information & Computer Sciences, Kanagawa Institute of Technology
‡Information Faculty, Kanagawa Institute of Technology
{s1385016@cce, s1185006@cce, ioroi@ic, h_tanaka@ic }.kanagawa-it.ac.jp

Abstract

This paper describes an indoor positioning system that uses ultrasonic signals and can cover a wide area. This indoor positioning system can be used in many practical applications, for example, detecting the location of moving objects, such as a person or a wheelchair, and navigation within a wide indoor area such as an underground shopping mall or a large hospital. The authors have previously proposed a solution for indoor positioning using ultrasonic sensors. However, the previous prototype had a shortcoming when it came to expanding the positioning area. An ultrasonic receiving unit that includes an amplifier, detection circuits and a one-chip microcomputer has now been developed and these units may be connected in a cascade style configuration, so making it easy to expand the positioning area. In a static test performed to evaluate the positioning accuracy, the offset positioning error was about 22mm and the standard deviation of the positioning results was about 2mm. In a dynamic test, 34 ultrasonic receiving units were installed in the ceiling and about 30.4m model railway track was laid on the floor, and the ability to identify the moving position of a model train on the railway track was confirmed. The proposed system configuration, its effectiveness and its accuracy were verified by these experiments.

Keywords: Position Information, Positioning, Ultrasonic Sensor, Information Infrastructure

1 Introduction

The global positioning system (GPS) is widely used for determining position in outdoor areas. It has become a universal system and the location information provided by GPS is widely used in navigation devices and many other service systems. However, for indoor positioning applications, no common system has yet been established and various possibilities are being investigated such as using wireless LAN systems. Position information is an indispensable requirement for the realization of "smart space".

Indoor positioning information can be made use of by many applications. The size of usage area and required positioning accuracy for some of these are shown in Fig.1. Some applications require a higher accuracy than can be satisfied by using wireless LAN systems. The authors have been investigating a positioning system that uses only ultrasonic sensors (without radio frequency transmissions) to satisfy accuracy requirements [1]. The previous prototype suffered a shortcoming in covering the positioning over a wide area due to the cable connections. The system configuration has been investigated and changed to use a simple cable connection which makes it easy to expand the area covered by the positioning system, hereafter called the positioning area. The positioning accuracy and the validity of the new system configuration have been evaluated by using a new prototype model and a model railroad train moving on tracks on the floor.

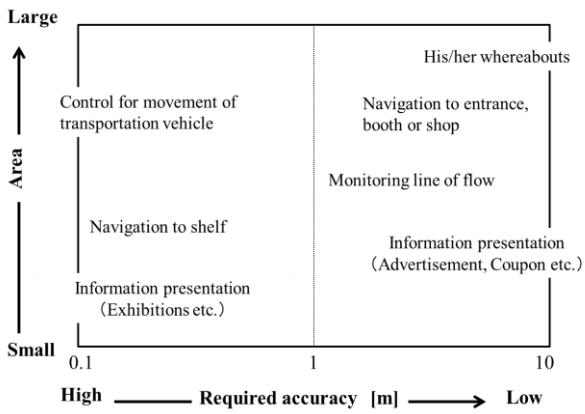


Figure 1. Application region and required accuracy

2 Positioning Principle and System Architecture

2.1 Positioning Principle

The system configuration involves the use of ultrasonic receiving sensors embedded in the ceiling and an ultrasonic transmitter attached to a moving object below. The basic positioning principle is shown in Fig.2. The clock of the transmitter unit is independent of the receiving part in the proposed system. Therefore, the propagation time cannot be obtained and only the relative delay times at each receiving sensor can be detected. Here, the delay time refers to the amount of time that has elapsed after the first receiving sensor detects the ultrasonic signal. Three delay times (t_1, t_2 and t_3) can be obtained from four receiving sensors as shown in Fig.2. The synchronization of all receiving sensors is maintained, since they are all connected to a PC which governs the timing, as described in section 2.2.

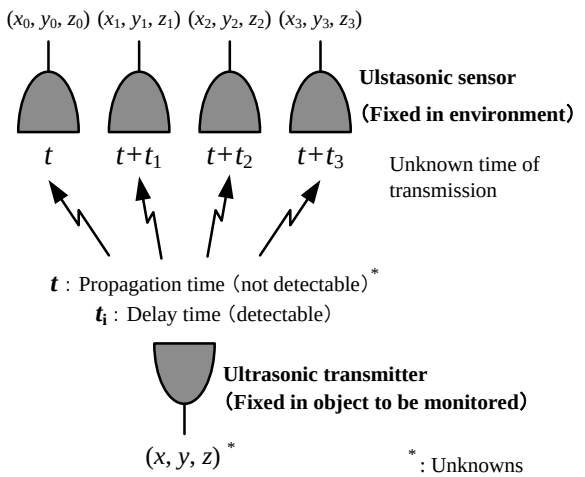


Figure 2. Basic positioning principle

The following mathematical equations can be derived by considering the distance and propagation time. Here, x_i, y_i, and z_i are the positions of the receiving sensors and are determined when they are embedded in the ceiling. The constant, c, is the ultrasonic signal velocity (dependent upon temperature). The variable, t, is the propagation time from the transmitter to the receiving sensor that is the first of all the receiving sensors to receive the ultrasonic signal (an unknown value). The variables t_1, t_2 and t_3 are the delay time at each receiving sensor (determined as described above). The variables x, y, and z are the coordinates giving the position of the moving object that is attached to the transmitter (these values need to be obtained). Since there are four unknown quantities (x, y, z and t), at least four receiving sensors are required in order to solve the following equations:

$$\begin{aligned}
&\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2} = ct \\
&\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} = c(t+t_1) \\
&\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2} = c(t+t_2) \\
&\sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2} = c(t+t_3)
\end{aligned} \tag{1}$$

2.2 System Architecture

The prototype system for lightening the workload of cable connection is shown in Fig.3. The upper diagram shows the previous prototype configuration [2] and the lower shows the new prototype. With the earlier configuration, the many cables from the detection must all be connected to the Receiver FPGA, which detects the timing of ultrasonic reception. As the positioning area expands, the load of cable laying works increases. The new configuration shown in the lower diagram uses RS-485, in which each unit is connected to others in cascade. This configuration results in a decrease in the work of running cables and makes it easy to expanding the positioning area.

In this prototype, the amplitude and detection circuit for the ultrasonic signal and a one-chip microcomputer (PIC16F1823), to capture the timing of the received ultrasonic signals, are implemented as a single receiving unit, which makes it easy to connect a number of units in a cascade configuration. An overview of the receiving unit is shown in Fig.4. The temperature sensor is used for compensation because the ultrasonic propagation velocity changes with temperature. In addition, the crystal oscillator (EPSON MA505) is used for accurate clock generation instead of using the oscillator implemented in the PIC microcomputer. The ultrasonic receiving sensor and transmitter are is Murata MA40S4R and MA40S4S, respectively.

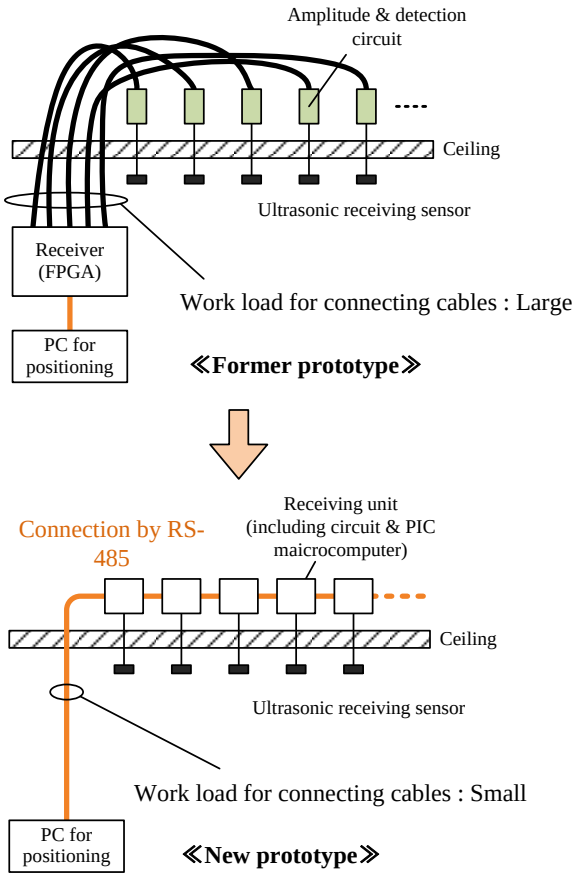


Figure 3. Configuration for expanding positioning area

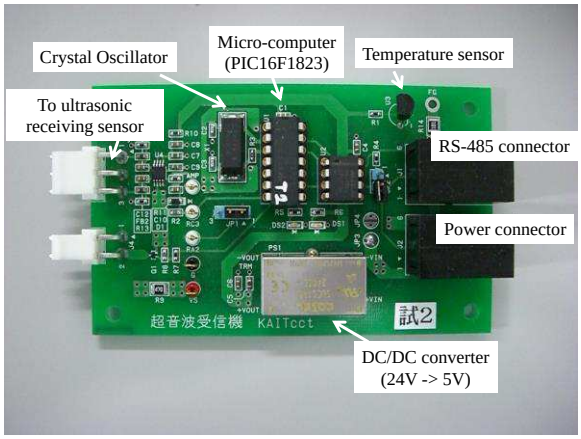


Figure 4. Overview of receiving unit

3 Evaluation Experiment

3.1 Static Test for Evaluation of Accuracy

The positioning accuracy of the prototype system was confirmed by experiment. The measurement target was set up on the floor (x=2300, y=1695) and its position was measured for accuracy evaluation. 100 measurements were carried out in order to evaluate variability as well as accuracy. The results are shown in Fig.5. The offset errors were about 23mm and 10mm for the x and y axis, respectively, and the standard deviation was about 2mm for both axes. This accuracy should be sufficient for the application shown in Fig.1.

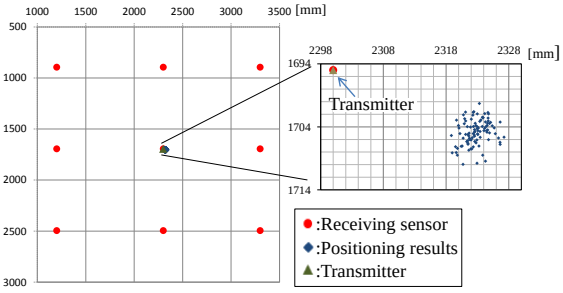


Figure 5 Results of static test for positioning accuracy

3.2 Dynamic Test for Moving Object

The practical usefulness of positioning system for a moving object was confirmed by experiment. Figure 6 shows a picture of experimental area. In this system, a model train, with an ultrasonic transmitter attached, was used as the moving object on a model railway track. The railway track formed a continuous loop with a total length of about 30.4 m. A total of 34 ultrasonic receiving sensors were installed in the ceiling. The height of the ceiling from the floor was 2.6 m. The experimental results obtained using this test configuration are shown in Fig.7, which illustrates the shape of the railway tracks, the positioning results and also the position of ultrasonic receiving sensors. These results were obtained as the train made two complete circuits of the railway track. Therefore, repeatability was also confirmed by this result. It was confirmed that the positioning error was within 100 mm in the xy plane. This should be sufficiently accurate to monitor the location of a moving person or, for example, a moving electrical wheelchair in an indoor area. Although abnormal numerical answers were obtained with a low probability (less than about 5%) in the positioning calculation, such results can be identified and omitted by taking account of the realistic position and velocity of the moving object and the propagation time of ultrasonic waves, etc.

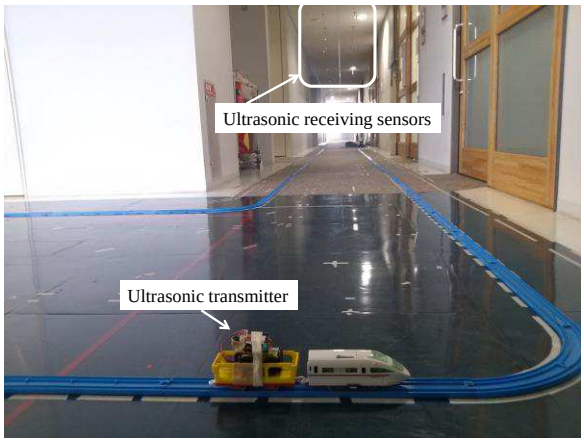


Figure 6. View of experimental layout

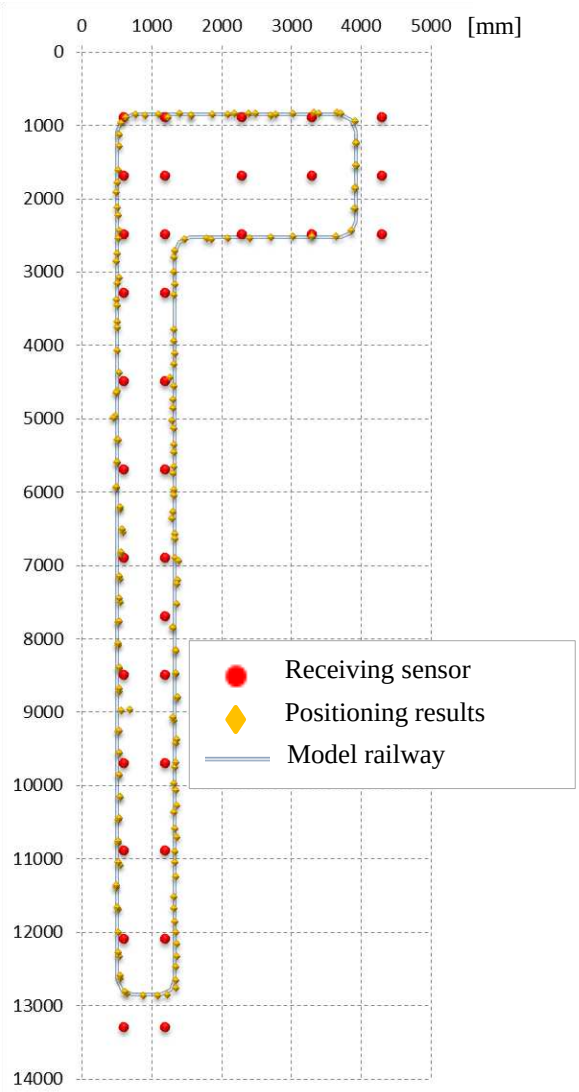


Figure 7. Result of dynamic test result on railway track

4 Conclusions

We have proposed a new architecture for an ultrasonic positioning system for indoor use over a wide area. The authors' developing ultrasonic positioning system requires neither synchronization between the transmitting and receiving units nor radio frequency signals. In this prototype, the amplitude and detection circuit for ultrasonic signal and a one-chip microcomputer (PIC16F1823) to capture the timing of the received ultrasonic signals are implemented as a single receiving unit so making it easy to connect a number of receivers in a cascade configuration. This configuration makes it easy to expand the positioning area. For an evaluation, receiving sensors were installed to cover an area of about 34m^2. In a static test carried out to confirm the positioning accuracy, the average offset error was less than 22mm and the standard deviation was about 2mm. This is considered to be sufficient accuracy for the assumed applications, such as navigation or position detection in an indoor area. In addition, a dynamic test for a moving object was carried out using a train on a model railway track. The accuracy in tracking a moving object was also confirmed by this experiment. The development and demonstration of an application system using the proposed positioning system is an area planned for further studies. The authors are now investigating the architecture of a cost effective positioning system which makes use of an inertial positioning using the acceleration sensor and gyro sensors built into smartphones.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 24500219.

References

[1] T.Hada et.al, Investigation and Demonstration of Local Positioning System using Ultrasonic Sensors for Wide Indoor Areas, Springer-Verlag, LNCS 6439:280-284, 2010
[2] H.Sunaga et.al, Discrimination of Multiple Objects and Expanding Positioning Area for Indoor Positioning Systems using Ultrasonic Sensors, Springer-Verlag, LNCS 7040:31-40, 2011

Identifying Physical Traits with Body-worn Accelerometers: A Preliminary Investigation

Akira Masuda[†] **Takuya Maekawa**[‡]
[†]School of Engineering
Osaka University, Japan
[‡]Graduate School of Information Science and Technology
Osaka University, Japan
masuda.akira@ise.eng.osaka-u.ac.jp, maekawa@ist.osaka-u.ac.jp

Abstract

This paper describes our preliminary investigation of the end user physical characteristics (e.g. gender, dominant hand, and skill at sport) that can be successfully estimated solely from body-worn accelerometers. For this purpose we use the huge quantities of data, which include 14880 labeled activities (e.g., walking and running) obtained from 61 subjects. We try to estimate various kinds of characteristics based on our simple idea 'When the activity sensor data of two users are similar, the physical characteristics of the two users may also be similar.' We consider that estimating the end user's physical characteristics will enable us to realize new kinds of applications that automatically recommend information/services to an end user according to her estimated physical characteristics such as gender and weight.

Keywords: Body-worn accelerometer, Physical characteristics, Daily activity

1 Introduction

Advances in sensing and wireless communication technologies have led to the low cost production of small wireless body-worn sensor devices. In the near future, end users will wear small sensor-embedded devices such as wristwatches, cellphones, and shoes, and their daily lives will be continuously recorded and recognized [1] by these sensors. Also, recently, several amusement parks (e.g., Disney World) have employed RFID sensor bracelets to provide personalized service for visitors. In this paper, we assume such a near-future sensor environment and attempt to estimate information about an end user's physical traits such as height, weight, gender, dominant hand, and age solely by employing acceleration data obtained from body-worn accelerometers. We believe that the automatic estimation of physical characteristics with body-worn sensors will prove very useful in many application domains. We show some examples.

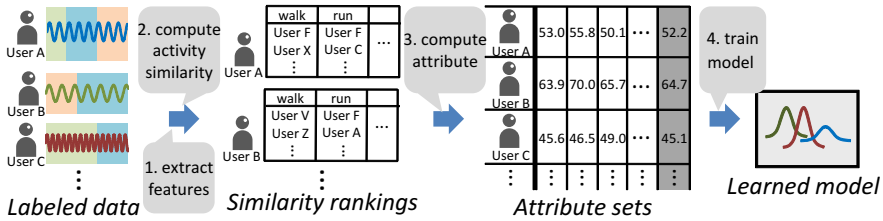
(1) We can construct a personalized recommender system that provides user-specific advertisement/information according to the user's estimated characteristics. For example, when the system wishes to provide an advertisement about clothes, it can automatically recommend clothes that match the user's height and weight. The system can also automatically recommend sports products according to the user's dominant hand. Furthermore, the system can effectively recommend health-enhancing products and provide the end user with tips about maintaining a healthy life according to her estimated health indicators, e.g., weight and BMI.

(2) While the investigation described below mainly focuses on known (apparent) physical characteristics such as height, dominant hand, and sports experience, we consider our method applicable to the automatic estimation of hidden physical characteristics such as diseases, health indicators, and sporting ability. In particular, the automatic diagnosis of diseases that affect body movements by using always-on accelerometers could be a significant application of wearable sensor systems. In fact, several physiology studies have determined the difference between normal subjects' gait acceleration signals and those of subjects with a disease, e.g., between healthy subjects and subjects with diabetes [2].

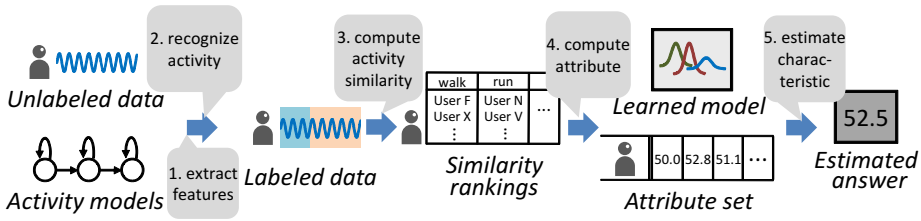
2 Estimation method

2.1 Outline of method

Our method consists of main two procedures; *Preparation* and *Physical characteristic estimation* as summarized in Fig. 1. In the *Prepa-*



(a) Preparation: Learn a model for estimating physical characteristics.



(b) Physical characteristic estimation: Estimate physical characteristics from unlabeled acceleration data.

Figure 1. Overview of the estimation method.

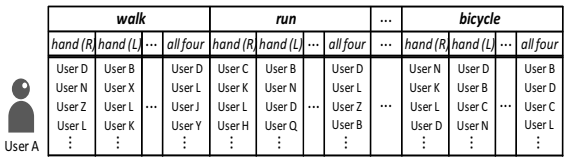


Figure 2. Example activity similarity rankings of user A. Each ranking includes training users in ascending order of similarity to user A. A ranking is computed for each activity class and each sensor variation.

ration procedure, we construct a model for each of the physical characteristics we want to learn/estimate in advance. A model estimates a given physical characteristic of an end user. In the *Physical characteristic estimation* procedure, we estimate a given physical characteristic of an end user by using her acceleration data labeled by the activity recognition system. We detail the two procedures below.

2.2 Preparation

We obtain labeled acceleration data including various kinds of activities from many training users in advance. Fig. 1(a) shows the outline of this procedure. In this procedure, (1) we extract features from the sensor data, (2) we compute similarities between the activities of training users according to each activity class by using the extracted features and construct rankings of each training user that include other training users in ascending order of similarity to the training user, (3) we compute attribute sets from the rankings to construct a model that estimates a certain kind of physical characteristic, e.g., gender or age, and (4) we then train the model by using the computed attributes and answers. An answer corresponds to the training user's physical characteristic value that we want to estimate, e.g., gender or age. We compute the attributes from the physical characteristics of other similar training users based on our idea that when the activity sensor data of two users are similar, the physical characteristics of the two users may also be similar. When we want to estimate a height, for example, we use the heights of similar training users to compute the attributes.

2.2.1 Computing similarities

We first construct a feature vector that concatenates the features extracted from sensor data obtained from a training user's accelerometers for each time slice. Then, we compute the activity similarity of each pair of training users by using their feature vector sequences that correspond to the activity class. We compute activity similarities between training users simply by using a Gaussian mixture model (GMM). Assume that we wish to compute an activity similarity between the 'walk' activities of training users A and B. We regard user A as a base user and user B as an object user, and compute the similarity between the object user's 'walk' activity sensor data and the base user's 'walk' activity model. We first model the 'walk' activity of base user A

Table 1. Activities performed in our experiment.

A	stand	D	sit	G	bicycle	J	use PC	M	play pingpong
B	walk	E	ascend stairs	H	brush teeth	K	draw on whiteboard	N	vacuum
C	run	F	descend stairs	I	wash dishes	L	write in notebook		

with the GMM by using feature vectors extracted from her 'walk' sensor data obtained from her accelerometers. We employ the EM algorithm to estimate the GMM parameters [3]. There are 5 mixtures in our implementation. Then, we compute the GMM likelihood of each feature vector extracted from the 'walk' sensor data obtained from the accelerometers of object user B. We assume that the average likelihood over the feature vectors corresponds to the similarity between the 'walk' activity of base user A and that of object user B. We compute the similarity for each pair of training users. By doing so, we can obtain the activity similarity between the 'walk' activities of a training user (base user) A and those of each other training user. That is, we can rank training users from the computed similarities. This ranking reflects the similarities of 'walk' activities to those of user A. The leftmost column in Fig. 2 shows an example of the ranking of 'walk' activities when we focus on the right hand.

2.2.2 Computing attributes and learning model

We construct a model that estimates a certain kind of physical characteristic by using the above rankings and the physical characteristics of training users. Here, assume that we construct a model that estimates an end user's gender. We focus on a training user A and compute a pair consisting of an attribute set and an answer. The answer corresponds to the gender of user A. We compute the attribute set by using the gender of similar users in the activity similarity rankings shown in Fig. 2. This is based on our idea that when the activity sensor data of two users are similar, the physical characteristics of the two users may also be similar. In each ranking, we compute attribute values by using the physical characteristics of the top-n similar training users. Here we employ a simple majority voting protocol. When n is 3 and the top-3 similar users consist of two males and one female, for example, we use 'male' as an attribute value. In our implementation, we use four n values (1, 3, 5, and 9), and compute an attribute for each n value

Table 2. Physical characteristic information used in our experiment.

name	value
dominant hand	{right, left}
gender	{male, female}
frequency of dish washing	{usually(3), sometimes(2), rarely(1), never(0)}
touch typing capability	{yes(2), somewhat(1), no(0)}
bicycle type	{upright, folding}

Table 3. Accuracies (percentages) of nominal physical characteristics.

	dominant hand	gender	bicycle type	frequency of dish washing	touch typing capability
naive	88.5	52.5	50.8	60.7	46.0
proposed	98.4	91.8	93.4	62.3	57.4

2.3 Physical characteristic estimation

Unlabeled acceleration data obtained from an end user are given. Fig. 1 (b) shows the outline of this procedure. In this procedure, (1) we extract features from the sensor data, (2) an activity recognition system [1] labels the feature vectors, (3) we compute activity similarities between the end user and each training user according to each activity class and each sensor variation by using the labeled data, and construct rankings of the end user that include training users in ascending order of similarity to the end user, (4) we compute an attribute set from the rankings, and (5) we then estimate a certain kind of physical characteristic by using its corresponding model learned in the *Preparation* procedure and the attribute set. The first, third, and fourth sub-procedures are identical to those in the *Preparation* procedure. In the fifth sub-procedure, by using the attribute set, the model learned in the *Preparation* procedure estimates the value of the end user's physical characteristics.

3 Experiment

3.1 Dataset

We used sensor data collected with wireless sensor nodes equipped with three-axis acceleration

sensors and sampling rates of 30 Hz. Each subject wore the sensor nodes on the wrists of both hands, the waist, and the right thigh. The subjects performed data collection sessions that included a random sequence of the activities listed in Table 1. Each subject also filled out a questionnaire that asked for information about the physical characteristics. We selected various kinds of physical characteristics related to the subjects ranging from basic information such as weight and gender to information about the activities listed in Table 1. The activity-related information included the dominant hand used in several activities, and the frequencies with which several activities were performed in their daily lives. The information about physical characteristics also included the types of objects used in the experiment, i.e., types of bicycle and types of vacuum cleaner. The characteristics used in our experiment are listed in Table 2.

3.2 Evaluation methodology

We evaluated our method using 'leave-one-subject-out' cross validation. That is, we regarded one subject as an end user and the remaining subjects as training users. We iterated the procedure so that each subject became an end user once, and we computed the estimation accuracies of the physical characteristics listed in Table 2.

We prepare naive methods for estimating physical characteristics and then compare them with our methods. A naive method for estimating nominal characteristics simply outputs the major nominal value among all training users. For example, when we want to estimate 'gender,' and the respective numbers of male and female training users are 10 and 20, the method outputs 'female.'

3.3 Results

Table 3 tabulates the classification accuracies of nominal physical characteristics for each method. The accuracy means the percentage of correctly classified instances (end users).

[Dominant hand] With 'dominant hand,' the accuracy of our method was very high, and better than those of the naive methods as shown in Table 3. The features computed from the 'draw on whiteboard' and 'write on notebook' activity similarities contributed greatly to estimating a subject's dominant hand when writing.

[Gender] We then focus on 'gender' results. We had considered it very difficult to estimate a subject's gender using only accelerometers. Contrary to our expectation, the accuracy of our method was good and much higher than the accuracy of the naive method. The accuracy was approximately 92%. Female's 'run' and 'walk' sensor data from the waist sensor were somewhat different from those of male's data.

[Type of object] As regards 'bicycle type,' the accuracy of our method was good as shown in Table 3. The features computed from the 'bicycle' activity similarities provided an important contribution. Because the wheel size of an upright bicycle is larger than that of a folding bicycle, the dominant frequency of the acceleration data obtained from the thigh when a subject rides an upright bicycle may be different from that for a folding bicycle.

[Daily activity related property] As regards 'frequency of dish washing,' the accuracy of our method was poor and it is not very different from that of the naive method. This may be because the problem is the four-class classification problem. Also, it may be very difficult to find difference between sensor data from subjects with 'usually' property and those with 'sometimes' property. With 'touch typing capability,' while the accuracy of our method was poor, it was somewhat better than that of the naive method.

Conclusion

This paper proposes a method for estimating physical characteristics by using body-worn accelerometers. As a part of future work, we plan to modify the method for calculating sensor data similarity. Also, we plan to estimate numerical characteristics such as height and weight.

Acknowledgement: We thank Dr.Yasushi Sakurai at NTT Communication Science Laboratories for supplying acceleration sensor data.

References

[1] L. Bao and S.S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive 2004*, pages 1–17, 2004.

[2] J. Petrofsky, S. Lee, and S. Bweir. Gait characteristics in people with type 2 diabetes mellitus. *European Journal of Applied Physiology*, 93(5):640–647, 2005.

[3] A.P. Dempster, N.M. Laird, D.B. Rubin, et al. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

From Linked Data to Linked Services: Developing Information Services for Sustainable Buildings

Vlado Stankovski **Boštjan Brank**
Department of Construction Informatics
Faculty of Civil and Geodetic Engineering
University of Ljubljana, Slovenia
{vlado.stankovski}@fgg.uni-lj.si

Abstract

In near future, it is expected that a large part of the existing building stock in Europe would need renovation and refurbishment in order to comply with energy efficiency and sustainability standards. The stakeholders (investors, architects, construction companies) need an adequate information support in order to make more informed decisions for specific renovation and refurbishment cases for which complex reasoning mechanisms are needed. The goal of the present work was to develop an architecture and implementation of an information system facilitating the publishing, searching and discovery of construction services over the World Wide Web (WWW) in a very refined way. The approach relies on Semantic Web technologies, the Web Ontology Language (OWL), the Resource Description Framework (RDF) and the Linked Data principles. It is shown that by using Linked Data and Services it is possible to incorporate complex information about the sustainability aspects of products and services, such as new Energy Efficiency (EE) approaches and Renewable Energy Sources (RES), which may then be applied to specific renovation and refurbishment cases.

Keywords: Linked Data, Linked Services

1 Introduction

It is expected that a large part of the existing building stock in Europe would need renovation and refurbishment in order to comply with energy efficiency and sustainability standards. For example, in Slovenia there are over 300.000 single family homes that would need renovation in the upcoming years. Hence, it is important to provide adequate services to the various investors in order to achieve greater energy efficiency and sustainability of these buildings. The stakeholders, such as building owners, architects, energy inspectors, materials, equipment and technology providers, designers need an adequate information support in order to be able to make more informed decisions in particular renovation and refurbishment cases.

The present approach aims at developing an information technology support for those activities related to energy efficient and renewable energy sources in buildings that are currently under way or will be introduced in near future. The complexity of information integration in this area arises from the fact that the various products and services are very diverse and multiple complex criteria must be used in their selection. An important goal of the present study was therefore to develop a model of an information system that can be used by the construction companies to offer materials, products and services for building renovation and refurbhishment over the World Wide Web (WWW).

Particular goals of the study were to develop:

- an appropriate architecture and a system prototype for the collection, processing and manipulation of data on building products and services;
- a model of products and services that takes into account the energy efficiency and sustainability aspects of the various materials, building products and services; and
- analytic capabilities, such as possibilities for comparisons based on building types and energy and sustainability indicators.

The first step in this process was to develop a model of a knowledge base of energy efficient and renewable energy sources, and buildings-related products and services. The approach also had to take into account European and Slovenian Law and guidelines for achieving sustainable buildings.

2 Related Works

Recently developed information technologies, such as OntoWiki [1] allow the conceptual modelling, creation and development of open and complex knowledge bases. The OntoWiki technology allows the use of the Web Ontology Language (OWL) [2] for conceptual modelling. Recently, the OWL has been used in various studies related to building sustainability and energy efficiency [3; 4; 5; 6; 7; 8].

A variety of exsiting technologies can be used for storing and quering RDF Schema (RDFS) metadata including, for example, Jena API, Sesame2, KAON2, Kowari Metastore and similar. These systems rely on traditional relational database capabilities and enable rich, highly detailed application programming interfaces for manipulating and accessing RDF data, as well as queries using different languages, including the SPARQL query language. Although the aforementioned approaches provide robust storage and query of RDFS data, there are a number of open issues related to security concerns, access control to metadata, transaction, replication, evolution of metadata, the strategy for the propagation of changes, and similar, which currently receive minimal attention. W3C's SPARQL Protocol is a very simple interface for Web services, which can be used to query RDF-databases. Although this technology has been intensively developed and tested, existing industrial applications do not fully exploit its capabilities for metadata management.

Recent Linked Data technologies [1; 8] promise an important shift in this domain. Most importantly, they allow for great flexibility and extensibility of the initial conceptual model of the knowledge base and for variety of intuitive browsing mechanisms. Based on an analysis of the end-user, application and system requirements, the OntoWiki technology [1] was selected for this study.

3 Building Energy Efficiency

According to some estimates 40% of the total final energy in the European Union (EU) is consumed by buildings [9]. This contributes to 40 % of total CO_2 emissions in the environment in Europe. Because this situation is similar in many other developed countries, it is a common goal to renovate and refurbish existing buildings in order to reduce the emissions in the environment and at least the unnecessary energy consumption. The EU has set up ambitious goals defined in a package of measures on climate-energy policy, labeled 20-20-20. The goals are to reduce Green House Gas (GHG) emissions by 20% (to the level of year 1990), to reduce primary energy use by increasing energy efficiency by 20%, and to increase the renewable sources in the primary energy balance by 20%, all until the year 2020 [10]. The Energy Performance of Buildings Directive (2002/91/EC, recast 2010/31/EU) (EPBD) prescribes stringent requirements for energy efficiency of new buildings and major renovations and a mandatory energy performance certificate for all buildings [10].

According to current estimates the cost-effective energy savings potential in buildings in the EU is 28%. The EPDB directive aims to ensure that after the year 2018 (an interim target is set by year 2015) all new public buildings have nearly zero energy consumption and by 2020 all new buildings have close to zero energy consumption. In several countries, such as Slovenia, these measures have taken up the form of action plans on energy efficient and renewable energy sources, which emphasise and require a sustainable and energy efficient construction and renovation of buildings [11; 12].

Hence, in the area of architecture, engineering and construction there is a strong need for a properly structured information knowledge base on energy in buildings, which would contain a large number of relevant information, and be accessible to the various stakeholders. An open knowledge base would allow for a proper assessment of the for low-energy new construction and renovation of buildings and disseminate knowledge for different user profiles (e.g., building owners, energy inspectors, equipment and technology providers, designers, manufacturers of energy certificates).

As part of the study, various information sources related to buildings were analysed. These included:

- information sources related to the building stock; and
- materials, products and services offered on the market.

Criteria for environmental, economic and social sustainability and energy efficiency of build-

ings are currently based on qualitative (checklists) and quantitative (Life-Cycle Assessment) approaches [13]. The European projects EIE BUDI, CA EPBD II and EPA ED proposed methodologies for quantitative estimation of the energy efficiency indicators of buildings based on structural and mechanical properties of the buildings. The idea of these emerging databases is to link them to the databases of the building stock and to obtain assessment of the energy performance for all buildings in Europe.

It is also necessary to use environment friendly construction materials with low embodied energy and low emission of pollutants, according to Environmental Product Declarations (EPD). Such energy efficiency indicators are expressed in quantitative terms and in the future will be publicly available as part of national databases of the building stock. Such databases will be upgraded by qualitative data on the various aspects, including environmental information. Some databases are currently under construction as a part of the implementation of the EPBD in Portugal and Ireland (EIE BUDI, CA EPBD II).

Because new products and services emerge on the market every day, it is necessary to use extensible and flexible technologies for the integration of such information. In the following section several use cases are discussed.

4 Use Cases

Our study further concentrated on an analysis of the motivation for creation of an open knowledge base of services for achieving energy efficient and sustainable buildings. In the course of the study, a number of experts were consulted, which led to the definition of use cases, which are listed in the following.

Use case 1: Publishing and searching for relevant materials for a specific case of renovation or refurbishment. Propose of the knowledge base is to be able to incorporate information on specific materials, systems, products, articles. These may include, for example, insulating materials, high insulation materials, organic insulating materials and their characteristics, energy efficient windows and properties, photovoltaics, heat pumps, heat recuperators, etc. The idae is that architects, designers and engineers will be able to access all information on all available materials and products in one place.

Use case 2: Publishing and finding the most effective services for the renovation of existing buildings based on requirements. It will be possible to obtain knowledge and information about known options and recommendations for refurbishment of existing buildings to achieve energy-efficiency in relation to their current characteristics (sample). It is understandable that not all methods of measure are appropriate for all existing buildings. This will enable prospective investors to choose between possible remedial measures for their building, according to his priorities (energy efficient, ecology, durability, economy, etc.).

Use case 3: Comparing the various products and services according to various sustainability criteria. The purpose of the knowledge base is to assist in analysis of the available alternatives of refurbishment according to the wishes and requirements of the investor, according to the certificate of the building. For example when one wants to renovate his building to become more efficient he has many possible approaches to choose.

5 Linked Data and Services - A Conceptual Model

When designing the knowledge base, in order to take care of the end-user perspective, a number of concepts were defined related to services and measures for:

- the improvement of the building's envelope, including addition or improvement of thermal insulation, placement and shading of energy efficient doors and windows, increase of thermal mass, optimum shape factor of the thermal envelope, etc.
- the reduction of the heating and cooling loads, including exploitation of principles of bioclimatic architecture, passive design, incorporation of passive heating and cooling techniques, i.e. cool coatings, control of solar gains, electrochromic glazing etc.
- the use of renewable energy sources, e.g. include solar thermal systems, buildings' integrated photovoltaic elements, hybrid systems, biomass, heat pumps etc.
- the use of intelligent energy management, i.e. advanced sensors, energy control (zone

heating and cooling) and monitoring systems;

- the improvement of the indoor comfort conditions in parallel with minimization of the energy requirements, i.e. optimization of the ventilation rate, use of mechanical ventilation with heat recovery, energy efficient systems of distribution, storing and heat emissions, energy efficient boilers, preparation of hot water, systems for cooling and air-conditioning, efficiency use of multifunctional equipment, i.e. integrated water heating with space cooling etc.

- the use of energy efficient appliances and compact fluorescent lighting etc. In the context of the transfer of measures that contribute to efficient energy use in public and residential buildings.

The developed ontology of the knowledge base is organised according to the types of services that may be offered by construction companies and renovation/refurbishment techniques. It includes possibilities for including comments from the building owners, energy inspectors, equipment and technology providers, designers, issuers of energy performance certificates and so on. Therefore, the developed ontology defines important entities related to energy efficient and renewable energy sources in buildings . On the basis of the developed ontological concepts, it is possible to describe very specific information: e.g. related to the selection of various ecological or energy-saving materials, process descriptions, measures and technologies, e.g. energy-efficient windows, properties of insulation materials, photovoltaic elements etc.

6 Architecture of the knowledge system and application

The identified requirements show the needs of the stakeholders for dynamic, well-structured and up-to-date access to complex information sources related to building sustainability. The key technology that was selected for the application development is the OntoWiki authoring and content-management system, an open-source product of the Lod2 project [14], which conforms fully with the Linked Data trend of the Semantic Web.

The system uses the Virtuoso advanced RDF storage system [15]. The efficiency of the RDF storage and the sophisticated knowledge-base access interfaces provided by OntoWiki API may be further used by applying a variety of reasoning technologies. The developed information system has a layered architecture including:

- User Interfaces Layer;
- Context Object;
- Application Layer; and
- Persistence Layer.

In the following, the layers of the architecture are described briefly in a bottom up fashion.

The *persistent storage* is based on an RDF storage adapter provided by OntoWiki. It supports authoring, versioning, Access Control Lists (ACL) and other necessary low-level functionalities. Moreover, it fully organises and optimises the content of the knowledge base for performance. Instances contained in the knowledge base concern various products, services, best practices, user management and authoring hierarchies, etc.

The *Application Layer*s main function is to contribute to a greater flexibility of the design of the end-user interfaces. A number of machine-accessible data interfaces, such as SPARQL [16] are supported, which makes the knowledge base suitable for future integration with other information systems in an open way. According to publication principles [17], OntoWiki makes all the resources accessible by their URI (provided that the resources URI is in the same namespace as is the OntoWiki instance). Furthermore, for all the resources used in the knowledge base additional triples can be fetched if the resource is dereferenceable. Another provided feature is that of a *semantic pingback*.

The *User Interfaces Layer* contains the exploration (end-user) interfaces for the knowledge base. They were developed by using the Cascading Style Sheets (CSS) Framework, the OntoWiki User Interfaces API and templates, and the RDF author. The exploration interfaces use the capabilities of the lower Application Layer of the knowledge base.

The knowledge base can be used to deliver various applications, for example, for the selection of best services for refurbishment of win-

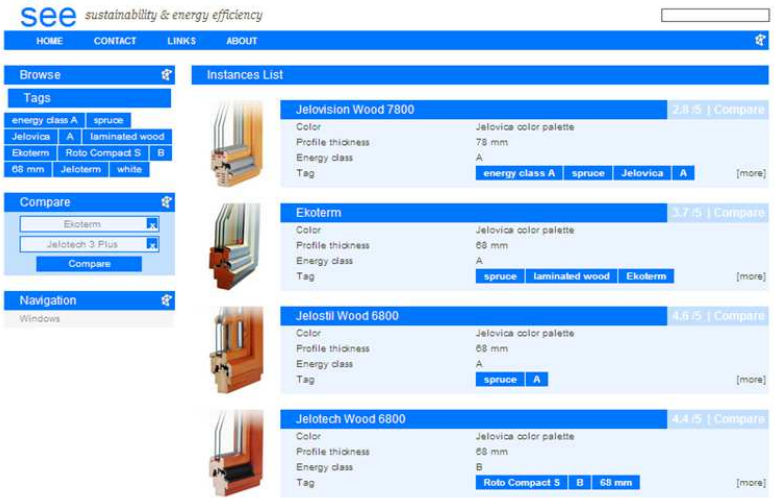


Figure 1. Improving the thermal envelope of the building through windows refurbishment

dows. An example of such application is presented in Figure 1.

Several advanced features are implemented and may be used, such as faceted browsing, content-specific browsing, full-text search, a visual query builder and a provenance-tracking mechanism. The first three mechanisms are sufficient for browsing and editing all the information contained in the knowledge base in a generic way, while the second two approaches can be used by more information-technology savvy users. Additional interfaces are designed to provide the publication of new resources, various comparisons, an analysis of refurbishment scenarios, tagging, commenting and rating, and other capabilities.

The knowledge system allows for the exploration of its content by using various mechanisms. The *Search* mechanism is designed to allow a free-text search in the knowledge base and the *Compare* mechanism can be used to compare two or more items, for example, two windows, two heating systems, etc. The *Select* mechanism makes it possible to select a list of necessary, appropriate products and/or professional services that the end-user has found in the knowledge base. The *Analyse scenario* mechanism makes it possible to analyse the overall effect of the selected materials and products in terms of the presentation of various sustainability and economic parameters. This helps the user when analysing various refurbishment or building scenarios.

7 Discussion and Conclusions

The purpose of this work is to present a model of an information system that can be used for spreading information on the energy efficiency and sustainability of buildings. The ontology developed in OWL proved instrumental in capturing the essenstial properties of construction services. It defines all relevant entities which are important when solving problems of building energy efficiency. On the basis of the developed ontological concepts it is possible to describe the specific materials, services and techniques (e.g. energy efficient windows and their properties, especially effective insulating materials, photovoltaic elements etc.) The ontology and the underlying knowledge base are designed having in mind the open-World assumption, which allows for its future extensions with other relevant concepts as and when they become available. The ontology can be used to annotate concrete specifications of materials, processes and technique, which are stored in a knowledge base. The use of the OWL language makes it possible to gradually extend the conceptual model with new knowledge items as and when they becomes available (e.g. on new materials, techniques and processes).

In near future one can expect rapid research and development in the field of sustainable assessment of buildings, in a way that environmental evaluation (based mostly on energy use and emissions of the implemented construction materials and products) shall be upgraded with eco-

nomic and social sustainability of the building. One of the scientific challenges in this respect is the comprehensive organization of the services and approaches and their effects on the energy balance and/or the building energy performance certificates. The energy related data will represent the core building information in the next years, being aware of the fact that energy and CO_2 reduction in buildings are European priority.

Acknowledgment

The operation part of this research is financed by the Slovenian Research Agency under grant ZeVS: J24262 Conceptual model and an open knowledge base on energy in buildings.

References

[1] Norman Heino, Sebastian Dietzold, Michael Martin, and Sren Auer. Developing semantic web applications with the ontowiki framework. *Networked Knowledge - Networked Media Studies in Computational Intelligence Volume*, 221:61–77, 2009.

[2] Owl. the web ontology language. `http://www.w3.org/TR/owlfeatures/`, 7 2013.

[3] Simone A. Ludwig. Comparison of a deductive database with a semantic web reasoning engine. *Knowledge-Based Systems*, 23(6):634642, 2010.

[4] Tracey Crosbie, Nashwan Dawood, and Saad Dawood. Improving the energy performance of the built environment: The potential of virtual collaborative life cycle tools. *Automation in Construction*, 20(2):205–216, 2011.

[5] Christina Diakaki, Evangelos Grigoroudis, Nikos Kabelis, Dionyssia Kolokotsa, Kostas Kalaitzakis, and George Stavrakakis. A multi-objective decision model for the improvement of energy efficiency in buildings. *Energy*, 35:5483–5496, 2010.

[6] Hans Erhorn, Tomasz Mroz, Ove Mrck, Fritz Schmidt, Lorenz Schoff, and Kirsten Engelund Thomsen. The energy concept adviser - a tool to improve energy efficiency in educational buildings. *Energy and Buildings*, 40(4):419–428, 2008.

[7] Pieter Pauwels, Davy Van Deursen, Ruben Verstraeten, Jos De Roo, Ronald De Meyer, Rik Van de Walle, and Jan Van Campenhout. A semantic rule checking environment for building performance checking. *Automation in Construction*, 20(5):506–518, 2010.

[8] Matija Konig, Jaka Dirnbek, and Vlado Stankovski. Architecture of an open knowledge base for sustainable buildings based on linked data technologies. *Automation in Construction*, page in press, 2013.

[9] 2010/31/EU. Eu direktiva o energetski učinkovitosti stavb, rs (2010). 2010.

[10] Marjana Šijanec Zavrl and Erik Potočar. Implementation of the epbd in slovenia. *Concerted Action EPBD National Reports*, 2011.

[11] Marjana Šijanec Zavrl. Energy efficiency and energy performance of buildings certificate. *Forum Media*, 2010.

[12] AN URE. National action plan on energy efficiency and action plan use of renewable energy sources 20082016. *RS*, 2008.

[13] Ignacio Zabalza Bribin, Alfonso Aranda Usn, and Sabina Scarpellini. Life cycle assessment in buildings: State of the art and simplified lca methodology as a complement for building certification. *Building and Environment*, 44(12):25102520, 2009.

[14] Sebastian Tramp, Timofey Ermilov, Philipp Frischmuth, and Sören Auer. Architecture of a distributed semantic social network. In *Federated Social Web Europe 2011, Berlin June 3rd-5th 2011*, 2011.

[15] Bernhard Haslhofer, Elaheh Momeni, Bernhard Schandl, and Stefan Zander. Europeana rdf store report. *Research Group Multimedia Information Systems. Faculty of Computer Science, University of Vienna,March 8, 2011*, pages 1 –30, 2011.

[16] Sparql. protokol sparql. `http://www.w3.org/TR/rdfsparqlprotocol/`, 3 2011.

[17] Jens Lehmann, Sören Auer, Lorenz Bühmann, and Sebastian Tramp. Class expression learning for ontology engineering. *Journal of Web Semantics*, 9:71 – 81, 2011.

The Development of Web Questionnaire with Business Process Modeling for Service Evaluation

Hisashi Masuda[†] **Wilfrid Utz**[‡]
[†]Japan Advanced Institute of Science and Technology, Japan
[‡]Research Group Knowledge Engineering, University of Vienna, Austria
masuda@jaist.ac.jp, wilfrid@dke.univie.ac.at

Abstract

In the servicizing economy, enterprises are required to have an improving cycle system including the customer side as well. But service evaluation that are not connected process is not interpreted to the structure of service evaluation and process. Therefore, we propose the integrated method of business process modeling and web questionnaire. In this here, we constructed two models for describing the differentiation of the service processes. The characteristics of first service process is a fixed process by definition. The other has a hybrid service process. By this method, we can have an improve cycle system including the evaluation from customer side.

Keywords: Service evaluation, Business process modeling, Web questionnaire

1 Introduction

1.1 Background

Every enterprises need to have an improving cycle system, e.g., PDCA(Plan/Do/Check/Answer) cycle in order to modify their business quality continuously. In the servicizing economy, every enterprises need to consider the service management as well. But the service process is very complex and not fixed. We need to record the change how to provide their service to the customers.

In addition, criteria of service evaluation have a variety by each person. For example, the evaluation criteria are diffrerent for the high maturity customer or not. And it is different from the situation and objective of customers. Therefore, service provider need to observe the evaluation from customers.

In this reason, we require the dynamic analysis in relaiton to both service provider and customer side. Previous our work, we found the two dimensions of satisfaction: novelty satisfaction and familiarity satisfaction from the point of dynamic aspect[5].

However we cannot understand the relation of service evaluation and service process. Because the method cannot be treated the differentiation of changing customer side and service provider side.

We cannot have the PDCA cycle if we cannot clear the structure regard to the dynamic aspect of such service evaluation. But, we can implement the PDCA cycle if we can understand the essential aspects.

1.2 Research Question and Objective

The criteria of service evaluation are different for each type of customer. And the criteria are affected by the change of characteristics of providing services. Therefore, we cannot interpret the which effect of service providers or customers if you cannot the understand the service evaluation with providing the actual service. However, current stage, it is difficult to understand the relation of the criteria of service evaluation and the service characteristics because we cannot get those data.

This research is proposed the solution using the integration method of business process modeling and web questionnaire with connecting service evaluation. Typically, to integrate the measure for the characteristics of customer side and the represented characteristics of service provider side, we can propose the interpretation between service evaluation and service process.

1.3 Approach

We construct the integration method of business process modeling and Web questionnaire for collecting the data from customer side. Because, we can propose the evaluation method for more concrete information to using the integrated method of service evaluation and service process. Like a

service, we need to this approach for the changing structure of service providers and customers.

We use the ADOxx® approach[9], that are developing OMI Project in university of Vienna as a developing platform, and the result of ADOxx approach is utilizing several business area[3]. The ADOxx is the multi-modeling language that main feature is business process modeling language. The high flexible extension is suitable for our integration method of process modeling and web questionnaire.

Our proposed method is to connect the questionnaire and each service process. Typically, in the case of fixed service process, we can ask the question without modifying the structure of business process. But, in the case of not fixed service process, we need to ask by the question based on the actual service process.

Section 2, we show the position of our research relative to previous works. Section 3, we explain the proposed method. Section 4, we constructed the prototype models for fast food case and Japanese Sushi restaurant case. Section 5, we explain the implication. Section 6 is conclusion.

2 Related Works

The direction of service evaluation research is to be changing to the dynamic aspect from the static one. The base research of customer satisfaction is the disconfirmation model [6] and SERVQUAL [1] which is related to service quality. The research focusing the dynamic aspect is frequent analysis in relation to time[7,8]. In addition, the research for dynamic aspect of customer satisfaction is the KANO model [4] developed by Kano.

On the other hand, in the business process modeling, there is an issue of describing the dynamic structure. It is the problem how to describe the process for the object that is not predefined. Typically it is the issue how to include the ad-hoc feature in the business process modeling.

Current stage, it is difficult to pursue to the dynamic aspect of service evaluation by business process modeling. Furthermore it is important to represent the dynamic aspect by current technique of business process modeling.

This research is proposed the connection between business process and the service evaluation. By this approach, we can represent the method for service evaluation from customer side by dynamic process. The following section, we explain our proposed method.

3 Method

We use the ADOxx approach as a developing platform. ADOxx is the business process modeling for server side. In this here, we provide to customers the developing platform for business process modeling. By this way, we can involve customers not only to the service evaluation but also to make the structure of service process.

We construct the integrated method of business process modeling and web questionnaire. User of this modeling tool login the our server from their tablet device. And input the service evaluation data based on the business process modeling. These data are collected the database.

We proposed two models for describing the differentiation of the service processes. The characteristics of first service process is fixed process by definition. The other has a hybrid service process. We call the former is the context-free service model and the latter is the context-dependent service model. Context-free service model do not design for adaptation of heterogeneity of customers. Context-dependent service model have a adaptation feature for each customer.

3.1 Context-free Service Model

Context-free service model is based on the fixed structure of business process. User of this model input their service evaluation with fixed the node and link. In this here, the questionnaire is asked by 5 scales questions to each customer. These items are customer satisfaction, expectation, the gap of expectation. The evaluation data is sent to the database.

3.2 Context-dependent Service Model

Context-dependent service model is based on the hybrid structure of business process. In this here, the model is pre-designed the default service process. After providing the service, service provider send the pattern of the service processes and decide the actual service process. Customer responds the modified questionnair.

4 Results

In this section, we constructed two prototype models. The first one is the context free ser-

vice model and the second is context dependent service model. Typically, Context-free service model for aiming to the static structure of service model is tested by the Fast food case. And Context-dependent service model for aiming to the hybrid structure of service model is tested by the Japanese sushi restaurant case. Japanese sushi restaurants have an ”Okonomi” style. Customers are the decision maker in relation to order process in this ”Okonomi” style. Therefore the service process is not be able to pre-design.

4.1 Case Study: Fast Food

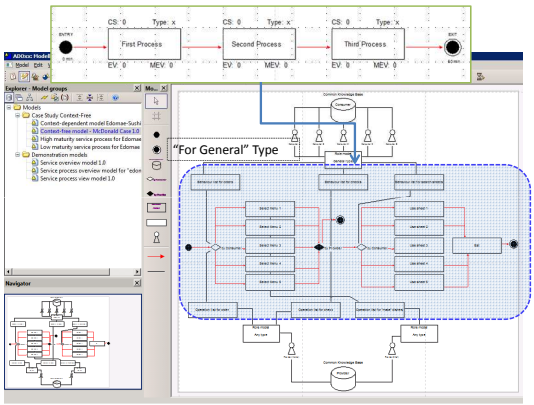


Figure 1. Fast Food Case

Fast food case (Fig.1) for operating all over the world is to realize scale-up by designing their service process for the all type of customers. This service type is enabling to us the fixed service process with web questionnaire to customers. User selects the property of process and input the data related to service evaluation. In this case, user doesn’t need the modified structure of service process.

4.2 Case Study: Edomae-sushi

In Japan, high-end sushi service restaurant (Fig.2) is required the knowledge in relation to the sushi service for customers [2]. The chef understands the customer knowledge by first impression of order process for drink. He modifies the service process based on this decision. By this way, we cannot pre-design the structure of the service process.

For this flexible service process, service provider decides the pattern of service process providing to their customers for the web questionnaire. And customers input their service

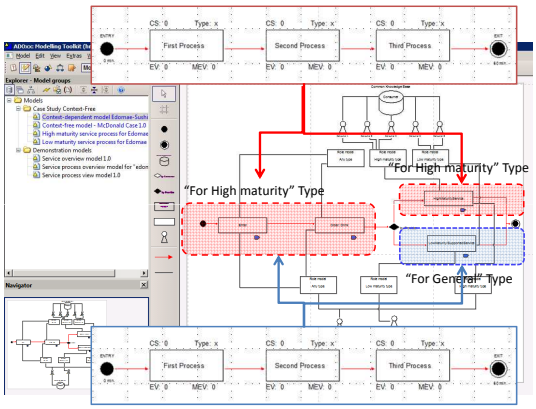


Figure 2. Japanese Sushi Restaurant Case

evaluation data based on the decided service process.

5 Implication

Implication for this research, we can interpret the service evaluation with connected the business process modeling and web questionnaire. By this way, we can implement the extended PDCA(Plan/Do/Check/Act) cycle for customer side.

In addition, like this questionnaire survey, we can divide the service process based on the maturity level of customer. To collect the data continuously, we are able to analyze the mechanism of changing process of customers’ criteria related to service evaluation.

6 Conclusion

In this paper, we proposed the integrated method of business process modeling and web questionnaire for deeper understanding of service evaluation. By this way, we can include the customer’s evaluation in the business process. Therefore, we can implement an extended PDCA(Plan/Do/Check/Act) cycle including the customer side.

However, in this stage, we constructed the prototype models only. Next step, we need to survey by actual data for empirical research. The targets of this empirical research are context-free service with static service process and context dependent service with flexible service process.

In this research, to develop the understanding to relation of service evaluation and service process, we can increase service quality by iterating improving process by PDCA cycle. Further-

more, in the point of interpretation of the service process, this research contributes to the knowledge engineering for the application to the service research.

Acknowledgment

We thank to Prof. Yoshinori Hara for advising our research. And we appreciate Prof. Dimitris Karagiannis supporting the OMI tool and technique. Empirical research of Japanese Sushi restaurant was mainly undertook by Dr. Yutaka Yamauchi. We appreciate useful comments from his research.

References

[1] A. Parasuraman, V. A. Zeithmal, and L. L. Berry,

A study on a generic development process for SOA design and implementation

Toshimi MUNEHIRA†

†Director R&D Department
OGIS-RI Co., Ltd, Japan
Munehira_Toshimi@ogis-ri.co.jp

Abstract

In order to optimize annual IT spending and to reduce the complexity of entire system architecture, SOA trials have been started. It is common knowledge that to design an SOA system we have to adopt the top-down approach, but in reality silo systems are being made, so these companies cannot reuse newly designed services, and cannot enjoy SOA's economic benefits. To prevent this situation, we designed a generic SOA development process referred to as the architecture of "mass customization."
To define the generic detail development processes, we did a case study on an imaginary company. Through the case study, we could define the practical development processes and found this could vastly reduce updating development costs.

Keywords: BPM, Generic Model, Mass Customization

1 Problem on a SOA development

It is common knowledge that to design SOA systems we have to adopt a top-down approach that begins with business modeling (Yoshida, Tanaka, and Une 2007[1], Dugan 2007[2], Yanagisawa and Mutoh 2008[3]).
It might be only in Japan, Japanese companies' SOA adapting cases accommodate this principle, but because of the limited business domain, they might be so-called "silo SOAs."
If these IT professionals continue building silo SOA systems, they will be sure to confront the problem that when interconnecting some SOA systems they find it very difficult because of different service definitions and master data, and as a result they cannot enjoy SOA's economic benefits.
To use BPM suites to implement SOA systems is one of easy methods, but we are afraid we would be easily locked in by the BPM vendor.
Originally, SOA allows us flexibility, but the SOA implementing tool, that is BPM suites, may decrease that flexibility.
To prevent such problems, we tried to develop a generic development process. In this report, we introduce the outline of this method and the results of validation of our SOA designing methodology

2 Requirements for a SOA

2.1 Differences between the SOA-oriented system and the legacy system

Table 1 shows the differences between the traditional system development approach and the SOA-oriented system development approach (Schmelzer 2007[6]). The orthodox systems are basically designed to last, but the SOA-oriented systems (SOA) are designed to change. This is the fundamental difference. Other items shown in Table 1 arise from this point.

Table 1.Differences between SOA and Legacy

Traditional Distributed Approach	Service Oriented Approach
Designed to last	Designed to change
Tightly Coupled	Loosely Coupled, Agile and Adaptive
Integrate Silos	Compose Services
Code Oriented	Metadata Oriented
Long development cycle	Interactive and iterative development
Middleware makes it work	Architecture makes it work
Favor Homogeneous Technology	Leverage Heterogeneous Technology

2.2 Requirements for SOA

2.2.1 Basic requirements

SOA was developed in response to users' need to adapt agilely to a change of business circumstances and to reduce system maintenance costs that might become a fixed cost, as well as to fulfill software engineers' desire to create flexible and stable system architecture (Nomura and Hara 2006[7]).
From the perspective of system development, we can find these common requirements for SOA.

1) Combine instead of coding
2) Immediately change systems according to a business change
3) Users do system modifications

2.3.2 Economical Requirements

Fig. 2 shows the expected cost model of SOA (Schmelzer 2007[6], Munehira and Shimada 2008[5]). One of the important reasons to choose SOA is to reduce maintenance costs. To realize this cost model is the most important requirement for an SOA development model.

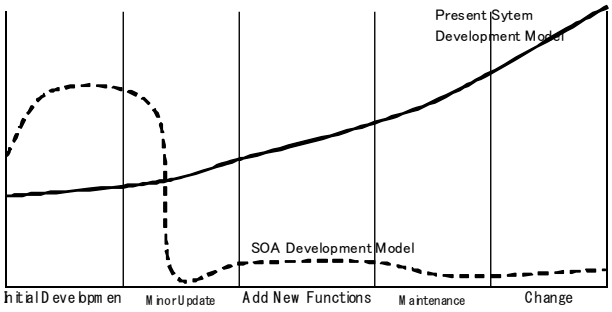


Fig.2 The SOA economics

3 Mass Customization as a reference model

Just as the waterfall development methods originated in the construction industry, we can find the similar solution of SOA in the manufacturing industry. That solution is based on "Mass Customization.

3.1 The definition of Mass Customization

Mass Customization is defined as follows (Pine, Victor, and Boynton 1993[8]).
"Mass customization calls for flexibility and quick responsiveness. In an ever-changing environment, people, processes, units, and technology reconfigure to give customers exactly what they want. Managers coordinate independent, capable individuals, and an efficient linkage system is crucial. Result: low-cost, high-quality, customized goods and services."
This definition has many similarities with those in Table 1.
In "mass customization," there are these four customization approaches (Gilmore and Pine 1997[9])

Collaborative: Collaborative customizers conduct a dialogue with individual customers to help them articulate their needs, to identify the precise offering that fulfills those needs, and to make customized products for them.
Adaptive: Adaptive customizers offer one standard, but customizable, product that is designed so that users can alter it themselves.
Cosmetic: Cosmetic customizers present a standard product differently to different customers.
Transparent: Transparent customizers provide individual customers with unique goods or services without letting them know explicitly that those products and services have been customized for them.

3.2 Requirements for Mass Customization in designing and manufacturing

Feitzinger and Lee pointed out three organizational-design principles for an effective mass-customization program (Feitzinger and Lee 1997[10]).

1) A product should be designed so it consists of independent modules that can be assembled into different forms of the product easily and inexpensively.
2) Manufacturing processes should be designed so that they consist of independent modules that can be moved or rearranged easily to support different distribution-network designs.
3) The supply network—the positioning of inventory and the location, number, and structure of manufacturing and distribution facilities—should be designed to provide two capabilities. First, it must be able to supply the basic product to the facilities performing the customization in a cost-effective manner. Second, it must have the flexibility and the

responsiveness to take individual customers' orders and deliver the finished, customized goods quickly.

3.3 SaaS, an example of Mass Customization

Salesforce.com typically shows the characteristics of Mass Customization (Munehira 2008[11]).

3.3.1 Customization approach

Salesforce.com adopts 4 customization approaches.

Collaborative: By introducing applied samples, it can clarify prospective customers' subconscious requirements. Since prospective customers can experience actual working systems, Salesforce.com can very effectively define their requirements.

Adaptive: After taking in a brief lecture, customers can make their own customizations themselves. If they need complicated customizations, such as an interconnection with their own systems, then an integrated development platform will be provided, which allows them to make their own programs.

Cosmetic: Users feel like they are using specially customized software, but Salesforce.com supplies only the usual service.

Transparent: As Salesforce.com adapts to various customers' needs, it continues updating systems without stopping the services. One day customers suddenly find that they can use new services without any configuration change.

3.3.2 Designing and manufacturing

Using software, Salesforce.com has built a required designing and manufacturing system. It can "turn its processes into modules and create an architecture for linking them that will permit them to integrate rapidly in the best combination or sequence required to tailor products or services." (Pine, Victor, and Boynton [8])

3.4 Requirements from Mass Customization principles

As the SaaS example shows, mass customization is exactly what we want to realize in SOA.
From the software engineering view, we translated Feitzinger and Lee's three principles into these two modularization requirements.

Modularized business processes and services
Design service systems as compounds of independent modules, and by combining these components, make it possible to provide each service with lower costs and fewer efforts.

Modularized service design and production process
In the design process, let business or system designers design the appropriate combinations of those individual modules, so that the business systems thus designed by business users are implemented as system services without any further business level modifications

4 The outline of a SOA development process

Fig. 3 shows a SOA development process model we designed to fulfill the requirements described in chapters 2 and 3.
The 1st step is business modeling. According to the company's strategy (BSC: Balanced Score Card), we design To-Be business processes that would achieve business goals.
In this business process modeling, we use the process model references to adjust the grain level of services (Munehira 2009[12]).
The 2nd step is service mapping. Activities in the system swim lanes are detailed into the BCE (Boundary, Control, and Entity) models. Controls are mapped to services stored in the service repository.
The 3rd step is service development. If proper services are not found in the service repository or mapped services lack some functionality, we will produce new services or modify the existing services. In this step, service modification is allowed but replication is strongly forbidden. Replication will seriously decrease the reusability of services.
From step 1 through to step 3, we do data designing using an orthodox method.
The 4th step is system implementation. Screens are designed simultaneously but independently. Designed business process data and related screens are set in a way that a process engine can understand. ESB-related configurations are also set.

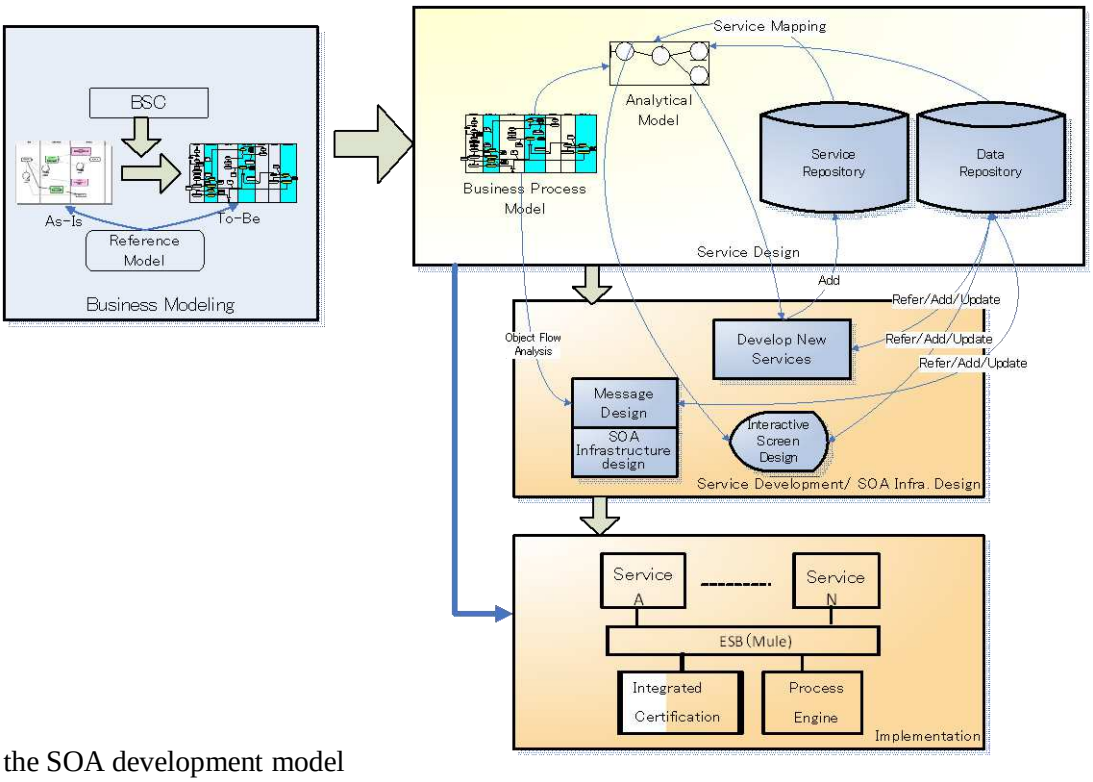


Fig.3 the SOA development model

5 Detail processes definition with a case study

In order to define the detail development processes and to evaluate the SOA economics, we did a case study on an imaginary company.

5.1 Preconditions of a case study

5.1.1 About an imaginary company

This imaginary company deals in sports bicycles and bicycle parts. This company has a wholesale department, nation-wide franchise shops, and a Web site, but it does not have any factory. This company imports bicycle parts from all over the world and sells them to bicycle fans through three channels, that is, specialty shops, franchise shops, and the Web site.

5.1.2 OSS-based development environment

The purpose of this study is to design a generic development process. For SOA, we have to design and implement both human processes and system processes. When we considered using COTS to build this environment, we found the following problems.

- It is a de facto standard to adopt BPEL for implementing a system process.
- Because BPEL engines are usually supplied from software vendors coupled with ESB, these engines usually have vendor original specifications. This would force us to implement special functions that work only under specific circumstances.
- BPEL for People is prepared for implementing human processes under BPEL conditions, but few engines are supplied, and it requires high-level technical skills. Therefore, it does not fulfill the requirements written in II-B.
- BPM suites support both processes. However, if we started to use a BPM suite, we would be locked in afterward by this tool.
- Many BPM suites require us to develop screens with this tool.
- The development process is also dependent on this tool.BPM suites support both processes. However, if we started to use a BPM suite, we would be locked in afterward by this tool.

In order to avoid the vendor-lock-in, we prepared OSS-based SOA circumstances for the implementation of this imaginary company's model

5.1.3 Development team

For SOA development, these three players are

required. We organized an international team.

- Business Modeler: Osaka
- SOA Designer: Tokyo
- SOA Implementer: Shanghai

5.2 Adopted development processes

5.2.1 Business process modeling

Fig. 4 is the To-Be process of "procure cycle parts designed to achieve the objective, "Shorten the lead time from sales order to purchase order." The upper activity diagram is the output of the 1st step, and the lower sub-activity diagram is the output of the 2nd step. Screen, Main-process, and DB correspond to BCE (Boundary, Control, and Entity).

The end of a business modeler's task is to register desirable services in a service repository. In Fig. 4, the 'T' mark shows this activity is mapped to some registered service. If a business modeler cannot find a proper service, then he or she creates a new service class in the "Service Repository package" and makes a mapping. After a business modeler finishes the business process modeling, a SOA designer's task begins.

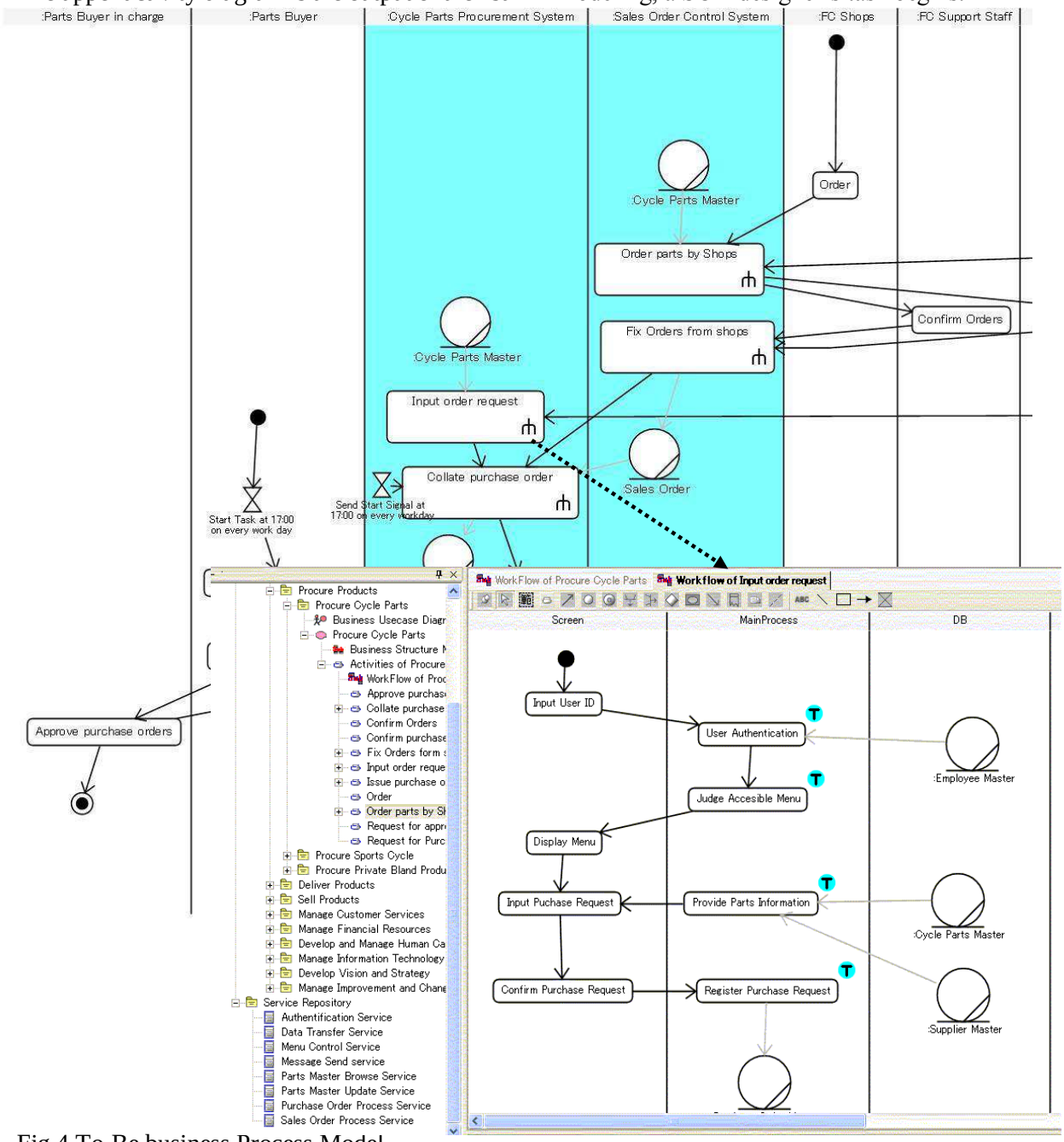


Fig.4 To-Be business Process Model

5.2.2 Service analysis and design

The SOA designer's tasks are defined as follows (Kranfzig, Banke, and Slama 2004[13], Oba et al. 2005[14]).

1. Service Analysis
    - Service candidate analysis
    - Conceptual model design
    - Service candidate refining
    - Service analysis
    - Message analysis
    - Service protocol analysis
2. Service design
    - Service design
    - Message design
    - Service definition documentation

Firstly we followed these steps, and modified them through the feedback discussion between team members.

5.3 Defined detail development processes

5.3.1 Detail process definition

Through this experience of SOA developments, we could define the generic and practical detail processes as below.

Service Analysis
- Service Candidate Analysis
- Service Candidate Optimization
- Service Analysis
- Message Analysis
- Business Process Analysis
- Service Protocol Analysis
- Feedback to Business Model

Service Design
- Message Design
- Service Design
- Process Design
- SOA Infrastructure Architecture Design

Service Implementation
- DB Logical Design
- DB Implementation
- Service Implementation & Unit Test

Application Design
- Screen Transition Design
- Screen Detail Design
- Application Architecture Design
- Test Design

Application Implementation
- Screen Implementation & Unit Test
- Executable Process Implementation
- SOA Infrastructure Implementation

Test
- Join Test

Fig. 5 (at the last page) shows the whole view and Tabel3 shows the sample of detail definitions

5.3.2 Definitions of a service category

We also found that services would be classified into these three categories.

1) Process Services
2) Business Services
3) Fundamental Services

Process services are almost equal to BPEL or workflows, and business services are called from process services. Fundamental services are security, ID management, access control, and so on.

Business services and fundamental services should be loosely coupled, independent, and reusable because they are called from various processes. However, process services are no t reusable. This policy is very important to guarantee the flexibility and agility.

This new knowledge is one of the important points of "Service Analysis."

We knew independence and reusability are CSFs for the SOA repository, but we did not know how to assure them.

Through this case study, we found this generic process would allow us to design proper services.

Table3 Sample of the definitions of the generic SOA development process

Activity	Subactivity	Details	Role	Input	Output
Service Analysis	Service Candidate Analysis	Extract service candidates from these elements −Services from the service repository −Activities from activity diagrams −Business Entities from activity diagrams	Service Analyst	Business Process Model	Service Candidates (Class Diagram)
	Service Candidate Optimization	Redefine service candidates based on relations that are CRUD, with business entities.	Service Analyst	Service Candidates (Class Diagram)	Optimized Service Candidates (Class Diagram)

6 Validation of economics of SOA

Table3 shows the difference between the FP (Function Points) based estimation (= Orthodox method) and the real data (= SOA development). By means of this comparison we found the following:

- Initial SOA development cost (1st Step) is almost equal to that of an orthodox method.
- In updating the system (2nd Step), SOA development cost is lower than that of the orthodox method as the Fig.2 suggests.

Table3 Actual results and estimation

	FP method	1st-Step	2nd-Step*
Estimation (Orthodox)	Data-Func Tran-Func No-Adjustment	36 69 105fpt	20+36* 37+60 57+96fpt
Translate to man days		141.13	179.18
Actual Results (SOA)	Analysis	10.00	20.00
	Design	30.75	
	Coding	46.55	49.10
	Test	49.40	55.60
	Total(man days)	136.70	124.70

7 Conclusion

SOA has been developed to meet users' needs to adjust smoothly to the changes of their business environment and to reduce system maintenance costs that might become a fixed cost, as well as to fulfill the software engineers' desire to create flexible and stable system architecture. The orthodox system development models cannot meet these needs.

The present situation in Japan with regard to SOA implementation methods is not good. Companies introducing SOA are making silo SOAs. If they continue in this manner, they will not achieve the purpose of SOA.

To prevent this situation, we designed a desirable model for introducing SOA, referring to the architecture of "mass customization."

To define the generic detail development processes, we did a case study on an imaginary company. In order to maintain independence from vendor software, we prepared the OSS-based SOA environment. Through this case study, we could define the practical detail development processes of SOA and found this would vastly reduce updating development costs.

References

[1] K. Yoshida, S. Tanaka, and Y. Une, "System development method based on EA and SOA", Toshiba review, Vol.62, No.9. 2007, pp72-75.

[2] L. Dugan, Use of SOA and web Services Technologies for EA Migration—Lessons Learned on How To Sort It All Out, Transformation and Innovation 2007.

[3] H. Yanagisawa and T. Muto, "Application of SOA in Software Product Development", Unisys Technology Review, Vol.93, August 2007, pp89-100.

[4] S. Latronico, F. Battista, How a business process vision may boost innovative ideas, Transformation and Innovation 2007.

[5] T. Munehira, and T. Shimada, Analysis and Procurement of Information System, JUSE Press, Japan, 2008.

[6] R. Schmelzer, Finding the Lowest Risk Path to SOA Adoption, Transformation and Innovation 2007.

[7] Y. Nomura, and H. Hara, "What is SOA? Now and Future," The Journal of the Institute of Electronics, Information and Communication Engineers, Vol.89, No.6, June 2006, pp506-510.

[8] B. J. Pine II, B. Victor, and A. C. Boynton, "Making Mass Customization Work", Harvard Business Review, September 1993.

[9] J. H. Gilmore, and B. J. Pine II, "Four Faces of Mass Customization", Harvard Business Review, January 1997.

[10] E. Feitzinger, and H. L. Lee, "Mass Customization at Hewlett-Packard: The Power of Postponement", Harvard Business Review, January 1997.

[11]T. Munehira, "A study on applying the mass-customization model on service businesses", 27th National conference of The Japan Society for Production Management, 2008.

[12]T. Munehira, "A Study for Refining a Business Modeling Method for SOA and SaaS Designing," APCIM2009 in Beijing, 2009.

[13] D. Kranfzig, K. Banke, and D. Slama, Enterprise SOA: Service-Oriented Architecture Best Practice, Prentice Hall, 2004.

[14] K. Oba, M. Hashimoto, and S. Fujikura et al., "The Status Quo and Challenges of Service-Oriented Architecture (SOA) Based Application Design," IPSJ SIG technical report, No.75, 2005, pp73-80.

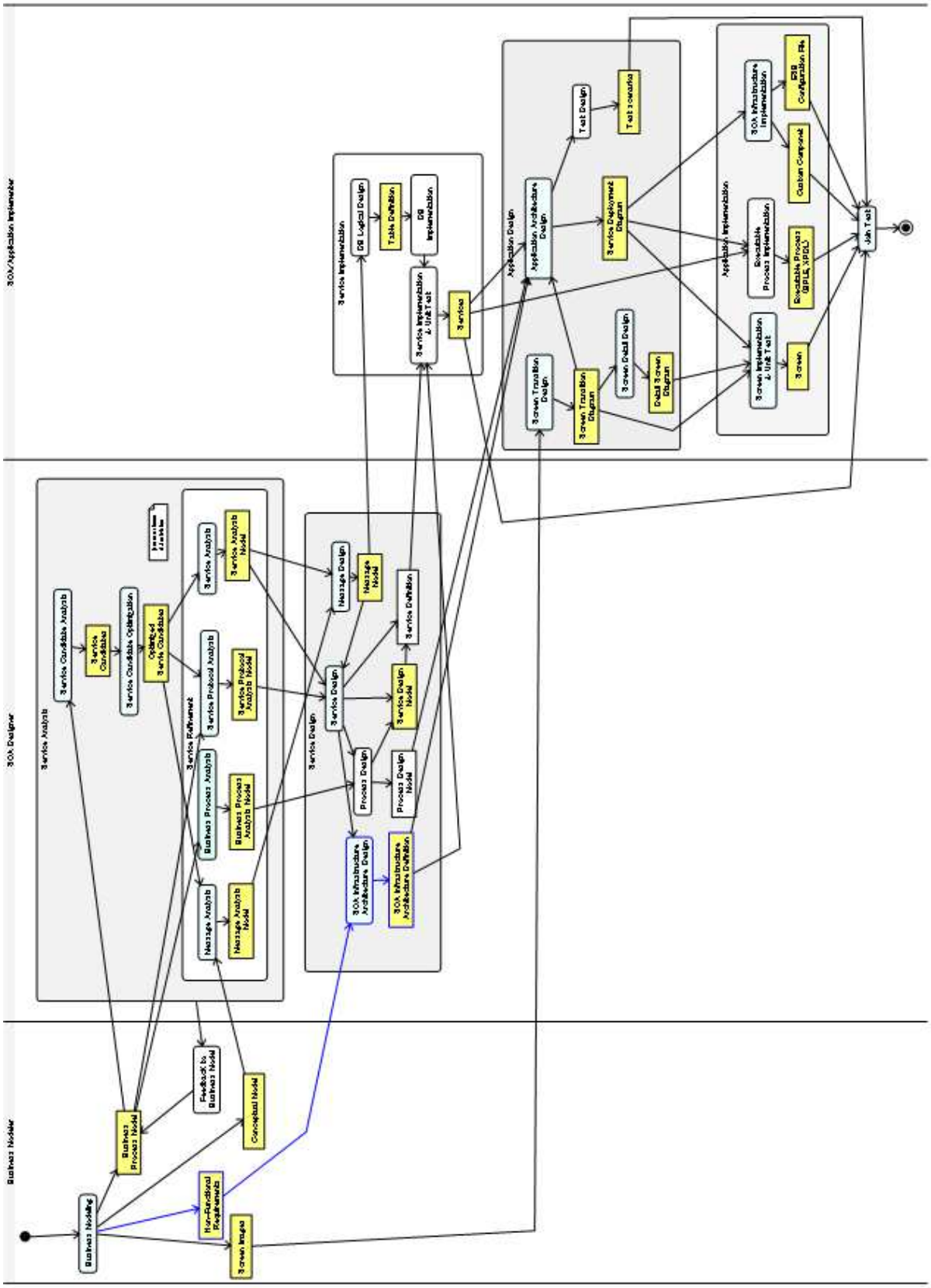


Fig.5 The Generic Development Processes

Continuous Improvement Approach of Software Process Based on Product Quality Inspection

Takahiro Seino[†] **Motoi Yamane**[‡] **Koichi Hayashi**[‡] **and Noriaki Izumi**[*]
[†]Faculty of International Social Studies
Maebashi Kyoai Gakuen College, Japan
[‡]Piecemeal Technology Inc., Japan
[*]National Institute of Advanced Industrial Science and Technology(AIST), Japan
seino@c.kyoai.ac.jp,
{motoi.yamane, koichi.hayashi}@pmtech.co.jp,
n.izumi@aist.go.jp

Abstract

We propose a methodology for the measurement and improvement of software quality according to the standardized processes and products based on a reference framework. In order to identify a root cause of poor quality, a real development project has been investigated from the standpoint of the software process. Through the detailed analysis with our proposed framework of traceability inspection, eight primary causes in three subject areas have been identified as patterns that result in poor quality. Furthermore, in order to mitigate the damaging effects on the poor quality, we suggest ways of reducing misunderstandings of the project's standards and references according to the identified patterns.

Keywords: Reference framework, development processes, quality control, requirement definition

1 Introduction

Generally, a project manager of development projects manages his/her project based on various indices for qualities, processes, effort, and so on. It is difficult to show whether products of the requirement analysis convey requirements correctly and exhaustively. Most projects carry out their own quality control, such as issue management. Using such traditional activities, to ensure quality does not only require additional effort, but also hardly predicts how much effort is needed.

The qualities which should be managed in such projects can be roughly divided into the following three aspects: whether the project proceeds according to the process standards, whether the products are created according to the product standard such as syntactical rules, and whether the documents convey users' requirements exactly and exhaustively. Traditionally, most projects have such standards, and measure only such surface-level indices, and expend most effort on corrections. However, efforts just increase and quality does not improve.

By exploring the root causes of these quality problems, we should be able to solve them by changing how the work is done. From the above viewpoint, this paper describes what project improvements can be obtained from comments in software reviews about the quality of products based on the software process, while analysing the quality subject which occurred to the product, based on the analysis results. We have demonstrated the proposed methodology by examining a real project as a case study. This project engages a reference architecture called AGUA[1; 2].

2 Reference Architecture for User-centered Development

In order to make an exact assessment of software product quality, a reference architecture is engaged as a set of project's standards and references for user-centered developments[1; 2]. The framework supposes that there are business experts who need to develop an information system, and ITC experts who can implement such a system. We call the former *user* and the latter *developer* in this paper.

The framework is designed with a glass box approach. The main concerns of users are that the system should provide functions and performances, not in technical details. On the other hand, for the developer, all required design infor-

mation needs to be disclosed. Therefore, AGUA provides the benefits of the black box for users, and the white box for developers simultaneously.

Process standard defines requirement analysis, design, implementation and administration processes. In this paper, we address issues and process improvements in the requirement analysis. As is known, the quality of the created products in requirement analysis is important and its poor quality carries high risks. If errors of requirements are discovered in the subsequent process, corrections will take time and effort compared with correcting in earlier process[3].

3 Requirement Definition Based on Reference Architecture

Requirement analysis process is the first process of a development, and its products are subsequent inputs. Investigating the current business activities consists of these four tasks:

Task 1: Drafting Products

Developer drafts a set of products based on the best explanations with all the information obtained from users, such as operational manuals.

Task 2: Interviewing Each Person in Charge

When drafts are completed, developers interview each individual person in charge. The purposes of these interviews are to check whether operating procedures and rules are correctly and exhaustively described in their products, and extraction of issues in the current operations. Developers keep track of interview.

Task 3: Modification products

Developers correct their products according to the comments of users. Developers perform internal reviews and proceed to the next task if the products pass the review.

Task 4: Inspection

Users check whether the products describe the current business activities correctly and exhaustively. Users also checks that there is a consistency between products. This task is performed in a meeting of formal inspections.

4 Analytic Experiment of Process Improvement Based on Product Quality

In order to investigate relationships between software process and product quality, a real project of a large-scale development is focused on, which engages AGUA framework. This project has over 4 subject areas and its budget is approximately 1.5 billion yen.

We describe the details of the requirement analysis process. We show the types of products created in the first subject area and their volumes in the bottom of Table 1. Now that workflow diagrams and business rule descriptions play the main roles in the requirement analysis, the quantity of them commensurately increases with business activities.

We have examined these products with an inspection, as a result we have 629 comments about quality issues of the above products.

4.1 Classification of the quality issues

We summarize the quality issues raised in the above observation. To make analyzing causes and planning improvements easier, we classify the issues as the following three aspects.

Process aspect issues are the difference between project reference and actual activities or products. We classify them into the following two sub-aspects. A *process accordance* is an issue with the lack of tasks to be performed or out of order. A *product accordance* is an issue with the lack of products to be produced on the basis of the references.

Format aspect issues are the non-conformance of syntactic aspects of the product standard. We classify them into the following four sub-aspects. *Format* is an issue where standard formats are not used. *Incomplete form* is an issue where missing inforation is required. *Traceability* is an issue where there is a lack of traceability information, or the information is wrong. *Style* is an issue where there is a violation of style guidelines.

Content aspect issues are misunderstanding the business activity explained by each person in charge. We classify them into the following four sub-aspects. *Wrong description* is an issue where the description about business activities differs from the explanation by the person in charge. *Incomplete description* is an issue where there is an incomplete description about business activities by the person in charge. *Granularity or consistency* is an issue that there are various granularity or inconsistencies with other products.

The result of this classification is shown in Table 1. It implies the following two points.

Table 1. Issues raised in our case study

	classification	Scope of Business Activities	A List of Actors	A List of Workflow Diagrams	Workflow Diagrams	A List of Business Rules	Business Rule Descriptions	A List of Business Forms	Glossary	A List of Problems	Records of Interview Mtgs.	Other comments	Subtotal	Total
process	process accordance												0	
	product accordance												0	0
formatting	format								1				1	
	incomplete form												0	
	traceability				35	1	41	3	4	1			85	
	style	5	6	3	13	2	3	2	1			2	37	123
contents	wrong desc.	24	1	3	172	8	72	10	4	7	6		307	
	incomplete desc.				19		4						23	
	granularity or consistency	3	1	2	123	2	29	2	2	4	8		176	506
	Total	32	8	8	362	13	149	17	12	12	14	2	629	629
	Quantities of products	1	1	1	37	1	65	1	1	1	32			

First, most issues are in the business process diagram and business rule descriptions. This is a natural result because these products concretely described the business activities explained by users.

Second, we need to pay attention to the distribution of issues. This is because there are no process aspect issues, the developer has complied with the process and product standards. Contrarily, there are many format aspect issues and there are still more content aspect issues. Content aspect issues shown in Table 1 are four or more times than format aspect. Through an examination of comments that are classified as content aspect issues, we conclude that most of the comments point out that the business activity was not described correctly.

5 Discussion

5.1 Analytic Investigation of Process and Quality

Each quality issue shown in Table 1 can be solved to correct each product; however this repair can improve quality only in the targeted subject. If we carry out the same process with another subject, the same kinds of issues will arise. In order to prevent these repetitions, we need a process improvement.

Developers need a certain time to learn the knowledge as well as have enough time to carry out the analysis. The developer gradually stores the domain knowledge through reading and understanding documents about the business activities and interviewing or observing the users. Similarly, they gradually acquire knowledge about standards.

We think that the reason why raising quality issues are important is because there are gaps in the knowledge between developers and users. We need to fill in these gaps, in order to understand correctly the information, which users give, to express them correctly in products and maintain consistency between documents. In order to acquire knowledge quickly and to disseminate it to the team, we need to improve techniques.

5.2 Process improvement and its effects

We add a preliminary interview task before the interview task of the requirement analysis. The new task aims at enabling developers to grasp the overview of business activities and fill in the gaps in their business knowledge. Therefore, this interviewee should be a business expert such as a manager.

Two other teams carried out the requirement analysis for the other two subject areas engaging the new process. In these teams, questions raised about quality issues on the inspections were significantly reduced. Moreover, we have observed the following three additional improvements.

First, both teams completed the task with less effort than the team engaged in the original process. This result has clearly shown up on schedule management. The original team had many delays, the longest one of which was ten days. Of course, we should take into account that each

Table 2. The number of quality issues before and behind the process improvement

	A	B
number of workflow diagrams	7	7
total of pages	25	22
number of issues	60	20
average issue / page	3.6	0.91

subject has its own difficulty and/or its own volumes, but in the subjects performed by both teams, most tasks were completed on schedule, the remaining being completed a few days behind or ahead of the schedule.

Second, the new process allows the team to understand the details of each business activity quickly. The main reason for this is that the team had grasped certain patterns peculiar to the business organization in the preliminary interview. This led to improving the quality of the draft version of products and narrowed down beforehand the questions about details of the individual business activity for the successive interviews.

Finally, the team mitigated the difference of notations, such as granularity or terminology, because they had introduced some notational rules. Through such efforts, the style of the product have became well-choreographed.

6 Related Works

The lessons learned from experiences with goal-oriented measurement have been structured into practical guidelines[4] for efficient and useful software measurement aimed at process improvement in industry.

As the following research, many efforts, in the Software Process Improvement (SPI) literature and an empirical study, have focused on the capability maturity model (CMM)[5].

In order to make the above research detailed in the software production activity, a systematic literature review has been performed[6] to identify and characterize evaluation strategies and measurements. Furthermore, as the latest research focused on real development, various metrics have been measured about time and quantity[7].

These efforts provide us with some lessons about partitioning business forms into development groups and determining staffing levels for each group. Unfortunately, the lessons are only concentrated on organizational aspects. Our proposed framework enables us to improve software process and activity focusing on product quality as real time feedback.

7 Conclusion

It is important for the large-scale development to improve the task of software process as well as the measurement and correction of the product quality. From the standpoint of quality driven SPI, we have proposed a methodology that enables a project to improve its software process through an indication about quality inspection. In order to remove the root cause in actual tasks, we provide the detailed aspects about quality inspections of products based on a software process of a reference architecture.

References

[1] Koichi Hayashi and Noriaki Izumi. Agua framework: Visualizing user-centered developments with a grass box approach in Japanese autonomy. In *6th Confenence of ISSJ (Information Systems Society of Japan) 2010*, 2010. In Japanese.

[2] A challenge to the amount and quality of information based on ontlogy. AIST Today. Volume 7, 2007. In Japanese.

[3] Barry W. Boehm. Software engineering. *IEEE Transactions on Computers*, 25(12):1226–1241, 1976.

[4] Lionel C. Briand, Christiane M. Differding, and H. Dieter Rombach. Practical guidelines for measurement-based process improvement. *Software Process: Improvement and Practice*, 2, 1996.

[5] Mark Paulk. *Capability Maturity Model for Software, Encyclopedia of Software Engineering.* John Wiley & Sons, 2002.

[6] M. Unterkalmsteiner, T. Gorschek, A.K.M.M. Islam, Chow Kian Cheng, R.B. Permadi, and R. Feldt. Evaluation and measurement of software process improvement? A systematic literature review. *IEEE Transactions on Software Engineering*, 38, 2012.

[7] U. Kohichi, H. Igaki, Y. Higo, and S. Kusumoto. A study of student experience metrics for software development PBL. In *2012 13th ACIS International Conference on SNPD*, pages 465 – 469, 2012.

User-Customizable Database Access Application on Smart Devices for Small Businesses

Daisuke Kikuchi **Toshiaki Ito** **Masaru Tezuka**

R&D department
Hitachi Solutions East Japan
{daisuke.kikuchi.hz, toshiaki.ito.tu, masaru.tezuka.fd}@hitachi-solutions.com

Abstract

In this paper, we propose DAAFS (Database Access Application For Small businesses), a user-customizable database access application for smart devices. With DAAFS, users can define both the user interface for smart device applications and the data access interface for the database servers, simply by editing XML files. As a native android application, DAAFS also has robust offline capability by default, and exhibits good performance under a variety of runtime conditions.

Keywords: User-customizable, smart phone, tablet PC, database access

1 Introduction

Modern smart devices enjoy a number of advantages over previous computing hardware, including user friendliness, versatility, availability, and low cost. These advantages have motivated an extraordinary shift in development toward smart mobile platforms and, more recently, their application to enterprise software systems. Unfortunately, while larger enterprises surge forward in the mobile space, most smaller businesses (i.e. those with fewer than 20 employees) have been deterred by the high cost of integrating new mobile clients with existing enterprise databases (for production, sales, inventory, etc.).

To lower this barrier to entry, we propose a customizable database access application for smart devices. The aim of this application is to allow users to define both the data interface of the smart device application and the connection interface of the database server using XML files. This should allow small businesses to make fuller use of mobile platforms in their enterprise systems, without the need for specialized programming or framework integration.

2 Requirements for Database Application of Smart Devices

We investigated a number of actual cases of mobile application development for small- and mid-sized businesses involved in housing inspection, manufacturing process management, shipping instruction, and so on.

All of the applications we examined ran on handheld terminals, i.e. specialized portable devices assigned to site workers and used to create, retrieve, update and delete (CRUD) data from proprietary databases in real-time. The common features of the user interfaces were (1) viewing data as list and (2) editing details of a selected record. The types of data varied from primitives such as strings and numbers to binary multimedia, such as digital photos and voice.

Our aim in each case was to replace the handheld terminals with smart devices, thereby cutting costs and increasing versatility. To this end, we established three distinct requirements for our smart device applications:

1) Balance among flexibility, ease of use, and cost of customization

Greater flexibility generally implies greater cost and greater complexity. Despite the greater flexibility of smart devices, our application must present a low-cost, easy-to-use substitute for handheld terminal applications.

2) Offline use capability

Since workers in the field or at production facilities often lack access to wireless networks, our mobile application must work offline as well as online.

3) Run-time performance

The computational performance of smart devices is still lower than that of PCs, but is generally better than that of handheld terminals. To

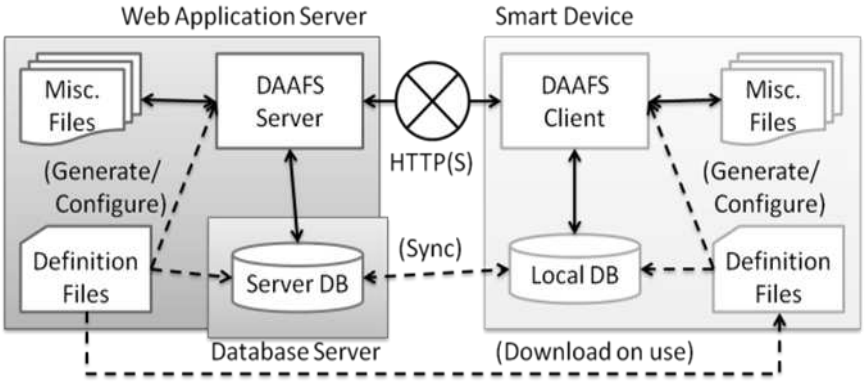


Figure 1. Structure of DAAFS

minimize user frustration, our application must preserve this performance advantage, meeting or exceeding the performance of handheld terminals across most functions.

3 Structure of DAAFS

Since our main target users are small businesses, we have chosen to name our application DAAFS ("Database Access Application For Small businesses"). Given the need to run both online and offline, DAAFS is implemented as a native smart platform (Android) application. A Web platform application was deemed unsuitable, as it could not be used in an offline environment, and would not perform as well as a natively compiled application.

3.1 Customization capability based on business case analysis

To find a good balance among flexibility, ease of use, and cost of customization, we examined the business cases of our target clients.

The essential functions of client applications in standard enterprise systems are encompassed by the CRUD (Create-Read-Update-Delete) methods on data. In our business cases, client devices were mostly used to view, store, and synchronize data managed in local clients and on server databases, so CRUD methods need to be easily callable from client applications and configurable for connection with any database, whether local or remote. Thus, we chose to limit application scope to basic CRUD functions.

3.2 System structure

The structure of DAAFS is shown in figure 1. Note that it is a multitier architecture [1] consisting of a "DAAFS client" and a "DAAFS server." The DAAFS client is a smart device application corresponding to the presentation layer. The DAAFS server corresponds to the data access layer, which connects to database servers in an enterprise system.

Each application has a definition file that includes a "screen definition" for generating client GUIs, and a "data definition," for configuring interfaces for data access. Customizing the application is a matter of creating and editing these definitions.

Note that, with proper customization, a wide variety of data (files, photos, etc.) should be accessible from and manageable through the application, in both offline and online environments.

3.3 DAAFS Client

The DAAFS client generates application GUIs for smart devices based on screen definitions. These GUIs allow a user to view and store data in enterprise databases. The GUIs are designed to support all basic CRUD functions.

The core aim of the DAAFS client is to establish a one-to-one correspondence between a table in the (local/remote) database and a screen to list data in the table. We consider this the most basic "data viewer" application.

We have also designed the DAAFS client to allow definition of what and how fields in a selected record are to be shown and edited, and what and how the GUI will transition between

these modes. These capabilities should cover the fundamental use cases of database-accessing applications in an enterprise system.

3.4 DAAFS Server

The DAAFS server provides common interfaces for hosted databases. Requests for authentication to connect with databases and commit CRUD transactions from DAAFS clients are executed via RESTful (Representational State Transfer) APIs [2], exposed via HTTP or HTTPS protocols. The data access interfaces are configured according to data definitions, which can be generated automatically from existing database schemas.

To support offline use, DAAFS includes a batch downloading function, which synchronize download multiple tables from a server to a local (client-side) database or vice versa. The definition file declares what data should be synchronized and how.

Definition files are stored on the DAAFS server and sent to the DAAFS client as necessary. Thus, the client can be automatically updated when data definitions change.

3.5 Advantages of DAAFS

We identify five main benefits of DAAFS for small enterprise users of smart devices:

1) DAAFS can access existing enterprise databases.
2) Users can customize DAAFS by themselves easily.
3) DAAFS has off-line use capability.
4) DAAFS maintains good run-time performance.
5) DAAFS can utilize smart device-specific features.

Those advantages improve usefulness of smart device applications for small businesses.

4 Evaluation of DAAFS

We implemented DAAFS and assessed its effectiveness in customizing database access applications. In this section, we discuss the platform and middleware DAAFS runs on, and provide examples of GUIs and data presented by the application. We also demonstrate its customizability.

Table 1. Middleware used in DAAFS

	Client	Server
OS	Android 4.0.3	Windows Server 2008 R2
Application Framework	Android SDK	ASP.NET MVC3 Framework
Database	SQLite3	SQL Server 2008

4.1 DAAFS platform and middleware

The DAAFS system is implemented on the middleware shown in table 1.

The DAAFS client is built as an Android application using the Android SDK [3]. Client-side data is stored in a SQLite [4] database for offline use.

The DAAFS server is constructed with ASP.NET MVC Framework [5] and connects with SQL Server databases for online use.

4.2 Example of Customization

Here we provide an example customization supporting the management of orders for some commodity. Figure 2 shows a sample data definition and figure 3 a sample screen definition. Note that the XML format is quite readable; most users should have little difficulty parsing and pattern matching this format for further customization.

Based on the data definition in figure 2, DAAFS finds a database with two tables, “Orders” and “Goods”. The attribute “id” of the <Table> and <Column> elements is the nominal identifier used by the database. The table “Orders” consists of five columns: “OrderID”, “GoodsName”, “Quantity”, “UnitPrice”, and “PickUpDate”. We can define the data types of each column using its “type” attribute. Available data types include “String”, “Integer”, “Float”, “Date”, “Photo” and similar. Note that these are abstract types defined within DAAFS, and will be cast to suitable types for the given database at runtime. Whether the value is a primary key or is nullable is determined by further attributes.

We also show screen images of the Android application dynamically generated with DAAFS client from the screen definition.

Based on the screen definition in figure 3, the client application generates two GUI screens for the “Orders” table: the list view screen shown in figure 4 and the detail view screen shown in fig-

```
<DataDefinition>
 <Table id="Orders">
  <Column id="OrderID"
        type="String"
        primaryKey="true"
        notNull="true" />
  <Column id="GoodsName"
        type="String"
        notNull="true" />
  <Column id="Quantity"
        type="Integer" />
  <Column id="UnitPrice"
        type="Float" />
  <Column id="PickUpDate"
        type="Date" />
 </Table>

 <Table id="Goods">
  ……
  <Column id="GoodsName"
        type="String”
        notNull="true" />
 </Table>
</DataDefinition>
```

Figure 2. Example of Data Definition

```
<ScreenDefinition>
   <Screen table="Orders">
    <ListView onSelect="Detail">
     <ViewItem column="OrderID" />
     <ViewItem column="GoodsName" />
     <ViewItem column="Quantity" />
    </ListView>

    <ListViewMenu>
     <MenuItem type=”Transit”>
      <TransitScreen table=”Goods”/>
     </MenuItem>
    </ListViewMenu>

    <DetailView>
     <ViewItem column="OrderID" />
     <ViewItem column="GoodsName"
            type="Reference">
      <Reference table="Goods"
            column="GoodsName" />
     </ViewItem>
     <ViewItem column="Quantity" />
     <ViewItem column="UnitPrice" />
     <ViewItem column="PickUpDate" />
    </DetailView>
   </Screen>

   <Screen table="Goods"> …… </Screen>
</ScreenDefinition>
```

Figure 3. Example of Screen Definition

ure 5. In DAAFS, each accessible table always corresponds to both a list view screen and a detail view screen.

The list view screen displays the list of records in the table. The columns to be displayed in the list are defined in the screen definition. In the example above, columns “OrderID”, “GoodsName” and “Quantity” are designated with <ViewItem> tags under the <ListView> grouping. Thus, the list view screen in figure 4 displays those three columns for the “Orders” table. The transition screen for when a record in the list view is selected is defined as the value of the “onSelect” attribute of the <ListView> tag. In figure 3, the value for this attribute is “Detail,” indicating that the detail view screen of the corresponding table should be displayed when a record is selected. Similarly, we can transit to another list view screen for the “Goods” table by selecting the appropriate menu option, defined under the <ListViewMenu> grouping.

The detail view screen shows the detail of the associated record. The columns to be displayed on the screen (“OrderID”, “GoodsName”, “Quantity”, “UnitPrice”, and “PickUpDate”) are defined with <ViewItem> tags under the <DetailView> grouping. The values displayed can be edited using native GUI components, such as a simple text box, a Date Picker, and so on. Suitable GUI components are displayed according to each data type.

We can also refer to values in other tables when editing. In this example, <Reference table=”Goods” column=”GoodsName”> is specified in <ViewItem>. As a result, when we press the reference button which has a magnifying glass, the values of the “GoodsName” column in the table “Goods” are listed for selection. Whether this column is editable can also be expressed in the screen definition.

All GUI screens are dynamically generated based on templates, so we do not need to specify further presentation details for screens that only differ in data displayed.

4.3 Ease of customization

As a point of comparison, we can estimate the

Figure 4. Generated List View Screen

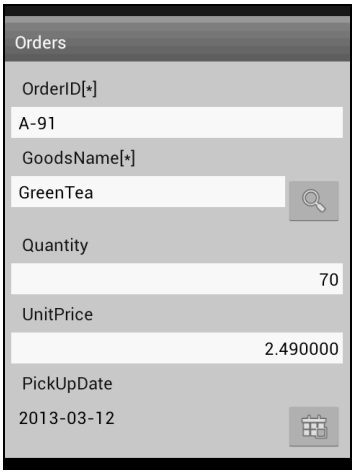


Figure 5. Generated Detail View Screen

Table 2. SLOC of Java Implementation Codes

Layer	SLOC
Presentation	3000 (Total of 15 screens, average 200 lines for each screen)
Data Access	3000 (2000 lines for connection with server database, and 1000 lines for connection with local database and data synchronization)
Total	6000

Table 3. SLOC of definition files

Layer	SLOC
Presentation	250 (Total of 5 tables, average 20 lines for each list view screen, and 30 lines for each detail view screen)
Data Access	75 (Total of 5 tables, average 15 lines for each table)
Total	325

SLOC (source lines of code) required to develop a new application equivalent to the customized DAAFS. Our estimate assumes the following:

- The application accesses one database consisting of five tables
- 10 columns for each table
- One list view screen and one detail view screen is equipped for each table
- Other five screens are provided to synchronize database, to configure settings, and so on

Conventionally, Android applications are implemented in Java. If a new application matching the above profile were developed from scratch, the scale of the project would be roughly as outlined in table 2, requiring 6000 lines total, 3000 for the presentation layer and 3000 for the data access layer. The reason this SLOC is so large is that we have to write presentation and data access code for each screen and table, even though their implementations are similar. In terms of developer time, it would likely take a month or more to deliver such an application.

Table 3 shows the SLOC for our DAAFS configuration files. Note that we can customize the application with only 325 lines total, 250 for the presentation layer and 75 for the data access layer. This represents only a few days of developer time.

5 Related Works

There are a number of frameworks supporting development of database access applications.

Microsoft offers Visual Studio LightSwitch[6] for rapid development of business applications that access various data sources, including SQLServer, SharePoint, and so on. The approach is almost same as in our framework, since LightSwitch uses an application definition to develop the user interfaces and data access interfaces of an application dynamically. LightSwitch offers high flexibility and ease of customization, but cannot leverage smart de-

vice-specific features such as local database and camera input.

Filemaker also offers a platform [7] for rapid development of iOS applications that access Filemaker and/or ODBC connectable databases. The applications are tuned to iOS and can leverage certain capability of iOS smart devices, and they can also run in offline environments by design. However, a Filemaker Pro server is needed to connect with existing online databases, significantly increasing development cost.

6 Conclusions

Our case analysis indicates that DAAFS would meet the needs of a wide range of small- to medium-sized businesses. Its ease of use, customizability, and robustness set it apart from existing solutions.

In future work, we plan to extend our work on DAAFS in several ways:

- *More exhaustive evaluation*: Though we have confirmed baseline reductions in application development cost, to show how effective DAAFS is in practical settings, we need to evaluate its applicability to small businesses solutions based on functional and non-functional requirements.
- *User-friendly tools for writing definition files*: We are currently developing a graphical screen editor, and plan to develop a data definition generator that makes use of database schemas or entity-relationship diagrams.
- *Adaptive GUI component selection*: We hope to improve the usability of the client by offering a larger selection of GUI components for supported data types, including multimedia types.
- *Handling business logic*: Simple use of CRUD methods through RESTful interfaces is not sufficient for modeling general business logic. We are currently investigating ways to express business logic in DAAFS using an XML format. This feature set may ultimately appear as an optional plug-in.

References

[1] Roy T. Fielding, Architectural styles and the design of network-based software architectures, University of California, 2000.

[2] Richardson, L, *RESTful web service*, O'Reilly Media, 2007.

[3] Android Developers, http://developer.android.com/

[4] SQLite, http://www.sqlite.org/

[5] ASP.NET MVC Framework, http://www.asp.net/mvc

[6] Visual Studio LightSwitch, http://msdn.microsoft.com/en-us/vstudio/ff796201

[7] Filemaker Platform for iOS, http://www.filemaker.com/solutions/ios/

Toward an Interactive Query Expansion: Concept driven and user feedback model

Krich Nasingkun‡, Boontawee Suntisrivaraporn‡*, Thepchai Supnithi*, Mitsuru Ikeda‡

‡School of Knowledge Science
Japan Advanced Institute of Science and Technology, Japan
‡*Sirinthorn International Institute of Technology, Thammasart University, Thailand
*Language and Semantic Technology Laboratory, National Electronics and Computer Technology Center, Thailand

Abstract

In the era of user-generated contents, a large different number of words are used for describing the same concept. Meanwhile, it is difficult to guess information we need using exact keywords. Query expansion is an approach to solve this problem by empowering a query with a possible relates information (terms or concepts). Existing approaches work well to establish all possible relevance information. But, personal interest is subjective to each person. Even same query from different users (or different search sessions) may intent to locate difference information. In this case, general relevance information from existing approaches cannot achieve better precision and recall. This paper proposes a design framework to make information searching process more effective, by using a combination of query expansion and user feedback. Live interest concepts are used to make a tailor made in the query expansion process. According to the iterative process of information searching.

Keywords: Query Expansion, Ontology, Live Interest, Web Search

1 Introduction

Searchers naturally prefer to post queries in their native languages. However, computer cannot understand these languages [1]. A simplified and straightforward way to formulate user queries is to use keywords instead of natural language, to approximate user's information needs. The majority of user queries carry a short form. For instance, the average number of terms per query is 2.4 for the Web users [2], and 1.59 for small-scale system users [3]. Very few web users adopt advanced searching options, e.g., Boolean operators, in query formulation [2] [3]. Users also tend not to refine the queries. However, high percentage of mistakes was also observed from users who occasionally employed advanced search options [2]. This behavior reflects the lack of subject knowledge of users to formulate more accessible queries. Therefore, users expect the search engines to completely satisfy their information needs based on minimal query input.

Results from this kind of behavior make huge irrelevant results for each searching process. Scrolling effects problem, the rank of web pages become significantly influences for Internet navigation. For example, the majority of search engine users do not look beyond the first three pages of search results [4] [5]. The order of search results presented by Google, etc. dramatically affects to the selection of the respective link; specifically, the subjects are significantly more likely to select the first and second suggested links. The links that presented from 8th to 10th are also likely to be chosen. However, those links presented from 3rd to 7th are very unlikely to be chosen. We can imply that, only a relatively small number of search results are exactly related to a subjective interest of users. Meanwhile, end user face difficulty to formulate keyword queries, especially short and unstructured queries. The question becomes, how to refine user queries into ones that are amenable to search services and perform the better suggestion about the context with semantic similarity of results.

Query expansion, is an approach to refine user queries in order to have more expressiveness, beyond the existing background knowledge of searcher. The aim of query expansion is to add new meaningful terms to the initial query. This process of adding terms can be manual, automatic or user-assisted process. New

suggest terms from term selection method should be provided, in the same context from an initial query. The contextual information can be acquired from many approaches such as relevance feedback, term co-occurrence, past queries, and knowledge models such as ontologies. In addition, to overcome a keyword search problem that make a huge irrelevant results and decrease scrolling effect, semantic search are used to implement on top of specific domain ontology. To provide semantically suggest terms for query process and the search service.

As described above, we have to deal with two problems: one is to refine the user queries, and another is to optimize the search service. By hoping that the search service returns reasonable results whatever the user input is. The focus of this paper is to refine the user queries instead of the latter one. We proposed a design of concept framework, by making use of a combination of characteristic of searching process and domain ontology approach to improve query expansion.

2 Related work

2.1 Relevance/Pseudo-relevance feedback

Relevance feedback is a fairly established technique for modification of the initial query using words from top-ranked or identified relevant documents [6]. It is an easier way to improve the retrieved document set as opposed to the user having to re-construct a new query. The main idea is to select important terms or expressions that appear in the documents that have been identified as relevant by the user. The refined query will help to retrieve results that move towards the relevant documents and away from the irrelevant ones.

The relevance feedback loop requires user to enter an initial query that results in a display of ranked documents (usually titles/abstracts). From this display, user makes relevance judgments and selects relevant documents. The relevant terms from these documents are added to initial query. The problem of this approach is, how to persuade users to provide feedback to results. In addition, whether it is relevant to the topic and/or relevant to the user.

An alternative to this is Pseudo-relevance feedback, where the top ranked n documents are assumed to be relevant. Terms from these documents are selected and used for expanding the query. Pseudo-relevance feedback methods improve retrieval performance on average. But, the results are not as good as relevance feedback. Problems arise when terms or phrases taken from assumed-to-be relevant documents that are actually non-relevant are added to the query. It causes a drift in the focus of the query. To tackle this issue, term co-occurrences is used to estimate word correlation for refining the set of documents in the query expansion [7]. An effective query sets by finding the potentially useful relationships among terms, called term association rules in text corpus are introduced [8]. Voorhees [9] found that assigning lower weights to added concepts could enhance retrieval accuracy. She used a factor between 0 and 1 for weighting added terms; f4modified formula [10] took into account the addition terms to the original query. She achieved the desired effect of bringing terms with low frequency to the top of the ranked list.

In conclusion, effectiveness of query expansion using relevance feedback can vary depending on many factors, such as choice of parameters in the term weighting process, number of relevant documents in the document collection, facilities provided for users to give good quality relevance feedback.

2.2 Past queries Technique

Past queries technique, Fitzpatrick and Dent [11] investigated the use of past queries to improve automatic query expansion, by using the results of past queries to form affinity pools in order to select expansion terms. The process works as follows: for a query that is to be expanded, up to three past queries that are highly similar to the current query are identified. The top 100 documents that were returned for each of these past queries are then merged to form the affinity pool. Running the original query against this pool identifies candidate expansion terms. Individual terms are then selected from the top-ranked documents using a Term Frequency-Invert Document Frequency (TF-IDF) [12] term-scoring algorithm. They demonstrate that this technique improves relative average precision for The Fifth Text REtrival Conference (TREC-5) collection by around 15%, from 21.4% to 24.5%. But, rather use result document from past queries as a terms selection source, Biller-

beck [13] propose to use expansion terms from past queries directly.

2.3 Corpus dependent

Most of the previous research concentrated on obtaining context from a document collection, using techniques such as stemming, clustering and term co-occurrence. During the late 60s and early 70s, many researchers investigated query expansion using term clustering. Similar documents are placed in a cluster. It is assumed that similar documents are relevant to the same requests. On this basis, if the query terms mapped onto one or more clusters the terms from the cluster(s) would be used for expanding the query. The usefulness of this technique was found to be marginal due to poor clusters resulting from small document collections. Another problem is that, a term can only belong to one cluster. For co-occurrence terms, since high frequency terms do not discriminate between relevant and non-relevant documents, the addition of these terms from query expansion is ineffective.

Lexical networks are another important source for deriving context. They contain domain specific vocabularies and relationships that have been automatically extracted from the collection. Text analysis tools can be used to develop the vocabulary for the lexical network. Lexical networks play an important role in the navigation process; Lexical relationships between terms are utilized to suggest additional terms. Some systems indicate the strength of the relationship between two terms. The general tendency is to disambiguate terms during the search process and not store the disambiguated terms prior to the search. Concept hierarchies can also be used to carry out query expansion techniques [14]. Instead of string based searches, concept based search are carried out which generally produce more useful results than the former [15].

Query expansion using corpus dependent knowledge models is more suitable for static document collections. For web collections, the knowledge models would have to be constantly updated and regenerated, because the collections on the web are more dynamic/fluid in nature.

2.4 Corpus independent (Ontology)

The problem with traditional relevance feedback techniques and corpus dependent query expansion is that they are content driven. The corpus content is analyzed to extract candidate terms for query expansion. This can only work if there are sufficient relevant documents to work with and also, these documents contain a reasonable set of terms that represent the subject area for the query. Corpus independent knowledge models do not suffer from this drawback.

Corpus independent knowledge models can be in the form of thesaurus or ontology. Ontologies provide consistent vocabularies and world representations necessary, to clarify communication within knowledge domains [16]. Ontology is more accurately described as “a classification, thesaurus or a set of concept clusters” [17].

Comparison of ontology-based query expansion techniques with relevance feedback techniques, the difference is that ontology based query expansion techniques will obtain term suggestions from the knowledge model, whereas relevance feedback techniques rely on having a reasonable set of relevant documents from which to suggest suitable terms. This is a point worth making because much debate has taken place on selecting the optimal sample size of relevant documents in relevance feedback techniques. With the ontology-based query expansion approach, sample size is not needed. Another similarity between the two sets of approaches is that the newly suggested terms still need to have weightings attached for the ranking algorithm. Finally the recent trend in collaborative filtering also bears similarity with ontology based query expansion, in the sense that the former is based on getting information from a community of users whose share the same interest as the searcher. Similar to ontology that usually be a ‘collective’ representation of a domain that has been derived from a community of domain-specialist users.

Domain specific ontologies are more suitable for work tasks. The terminology in these ontologies is less ambiguous; therefore short queries can be expanded with a higher chance of accuracy. General ontologies would be suitable for information type broad queries, however the query expansion process may need some guidance or interaction from the user. The trend is to use multiple ontologies to satisfy a user search query [17]. To improve precision and recall in ontology approached, Chauhan [18] propose a framework that use Domain ontology with

mathematical model, which control by specific threshold to compute similarity of concept, to maintain recall. The semantic distance, layer factor and degree of upper concepts are computed. The expanded term only be added if it consists of semantic similarity within a threshold. Thus, the relevancy of results is improved; as the expanded terms are closely related to seed query concepts.

However, ontologies have their problems. The first issue is related to vocabulary mismatch between the query terms and the concepts in the ontology. A mapping process needs to take place to overcome this problem. Secondly, if an ontology for a particular domain does not exist then a lot of effort is required to construct ontologies from scratch not just from a technical point of view but more important the process of knowledge extraction from domain experts and arriving at a consensus view [19].

In conclusion, existing approaches are work well to collect all possible interest terms or concepts. The remains problem is that, how to distinguish irrelevant interest with the relevant ones. Because, interesting is subjective to each person. Even for the same seed query from different user (or different search session), may be intended to locate difference information. According to Chauhan [18] proposed, if we can find the right threshold then we may success to improve precision and recall. But it's no generic value for everybody. Instead of using generic threshold value, this research will propose information search framework, by using user live interest as a condition to apply query expansion. To distinguish irrelevant and relevant one, by using the ontology approach and user feedback information. Output of this framework will help to make a subjective set of user live interest, that efficient to maintain precision and recall along information searching process.

3 Proposed Architecture

In information searching process, people always use first query as seed information to examine the reality of existing information. User mental model about interested information being develop during information search process. Not just only final interest result, Iterative process of information searching also provides an experience to searcher, which may effect to their information need. Relevance and irrelevance information are learned during search process, users always move towards the relevant and away from the irrelevant ones. The proposed architecture will provide mechanism to reach the interest information by keeping concentrate on user live interest. This framework allow user to experience and decide their interest that subjective to what they want.

3.1 System Architecture

The proposed architecture of semantic query expansion (Figure1) is comprised of 4 main modules: Query Analysis, Semantic Query Expansion, Semantic Inferencing and Result and Feedback. Three main support components: General ontology (WordNet [20]) and Domain specific ontology are used to identify the concept that relate to seed query and Concepts storage. Concepts storage is a database to store the whole relevance concept that already identified for each search session.

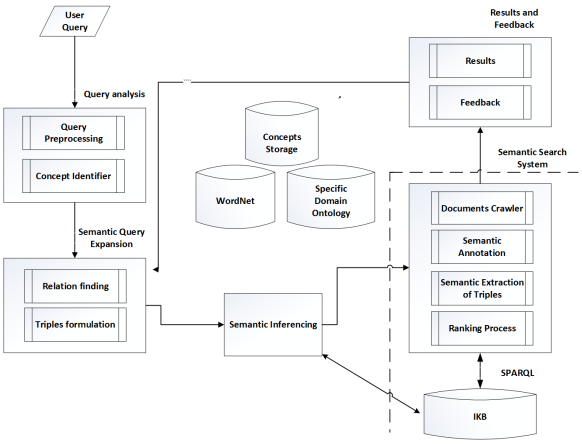


Figure 1. System Architecture

After feedback results are identified, relevant and irrelevant concepts are computed. All concepts are stored in Concept storage. Unused concepts will not consider in the next iterative process in the same search session, even it may rise during the next iterative process.

Information Knowledge Based (IKB) stores RDF triples that are output from web crawler, semantic annotation and instance extraction process. However, Semantic Search System that provide semantic results from World Wild Web to complete the propose system is beyond the scope of this paper.

By this proposed framework, users can get support information for all possible relate

concepts (by ontology approach), which help to narrow down to the interested information (by user feedback). This framework also allows users to learn and adjust mental model according to results from each iterative loop. Finally they can get information they want, not only related documents.

3.2 Query Analysis

Query Preprocessing: Seed (first) query from user be cleaned and examine to fix a possible miss spelling, while remove stop word and stem the terms.

Concept Identifier: The assumption in this modules is, Searcher may not be clear what information they looking for, even they can state some concept in first query. By use General domain (WordNet) and Specific Domain ontology as an input to identify the relevant concepts, such as upper concepts, lower concepts and synonyms concepts.

3.3 Semantic Query Expansion

Relation finding: After identify the concepts that may relate to the searcher interest. This module will discover the possible relation around the concepts by using information from domain specific ontology. Domains and ranges of relations are taken to be considered.

Triples formulation: all possible triples are formulated, according to previous modules.

3.4 Semantic Inferencing

By consider the relation in each triples, this modules will perform the relevance semantic rules to find implicit information in Information Knowledge Base (IKB).

3.5 Semantic Search System

This system response to create and maintain IKB, by performing a web crawler, semantic annotation and semantic extract information into RDF triples. The importance role of this system is to locate the relevance information according to semantic queries from the previous modules.

3.6 Results and Feedback

Results: All relevance results will show to searcher. If results are subjective to their search intention then they could finish.

Feedback: If searchers do not get information that they want, then navigation to each result will be used as feedback to system. Searcher can continue search process by traverse to the most relevance information. The proposed system will use this feedback, as a condition to state relevance and irrelevance concept for next iterative loop.

All concepts in the previous step will be kept in Concepts storage to use to identify relevance or irrelevance concepts from user interest (from feedback module). By this way, searchers can experience with all relevance information, while away from irrelevance ones.

4 Use case scenario

In this session we will describes use case scenario of the proposed system. To make understanding of how it work. Tourism Domain is used as specific domain ontology.

4.1 Users query

Example seed query submit is "*Hotel near JAIST*". From this seed query, it may not clear what exactly he wants. From human senses, this query can be related to a huge of information. For example, he may want to book hotel, interested in a sign seeing, or interested in any kind of accommodation near JAIST. His intention is subjective and may change during the search process.

4.2 Query analysis

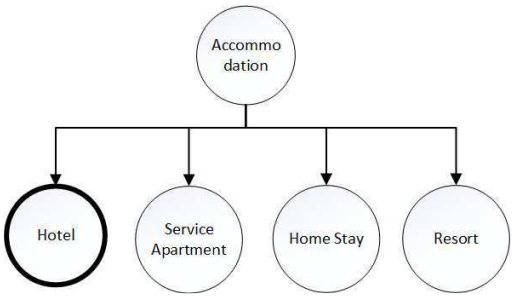


Figure 2 Accommodation Concepts

After preprocessing to clean and identified the relevance concepts, Figure 2 show the root concepts and relevance structure according to knowledge from Tourism domain ontology. The rest of keywords also need to identify the relevance concepts too.

Hotel -> Accommodation, Service Apartment, Home Stay, Resort

near -> near, close to, walking distance to, next to

JAIST -> Attraction, Location, University

From this set of concepts, we may guess that searcher may just want an accommodation. It is possible to be any kind of accommodation. In this case, an instance of concepts are specified ("*JAIST*"), that mean searcher have already clear in mind about target his want. In this case, we can use it as an instance value of all possible concepts.

4.3 Semantic Query Expansion

To formulate all possible semantic triples, combination of relevance concepts are used. By consider the domain ontology structure (relation, properties, domains, ranges).

(Hotel, near, Attraction ="JAIST")
(Hotel, close to, Attraction ="JAIST")
(Hotel, walking distance to, Attraction ="JAIST")
(Hotel, next to, Attraction ="JAIST")
…

(Resort, near, University ="JAIST")

4.4 Semantic Referencing

In specific domain ontology, domain experts will also design a relevance rules in order to inference a new knowledge from existing facts.

X near Y, Y near Z → X near Z……(1)
X near Y, Y next to Z → X near Z……(2)

In this example assume that 2 rules that relate to relation "near" are defined as (1) and (2). This rules are used to inference to find other possible (implicit) triples that may exist.

4.5 Semantic Search

Semantic search will compute the matching triples from semantic queries with both information in IKB and implicit information from the previous process. Example output as follows:

…
(JP Hotel, walking distance to, JAIST)
(Tsurugi Apartment, near, Tsurugi station)
….

4.6 Results and Feedback

In this module, all semantically related results are provided according to relation of concepts in ontology structure. Searcher may satisfy information they looking for or they have to select which one is the most related to their interested. Feedback module will help to make clearer about information their looking for. After relevance ones are selected, an unused concept (upper and sibling concepts) will be marked as irrelevance and never take to consider in the next iterative loop. System will assume that, the selected ones are only related to user interest. Relevance concepts are use as the starting point for next iterative loop.

In this example, if "*Tsurugi Apartment near Tsurugi station*" is selected. After that, all other accommodation concepts (Hotel, Home stay, Resort) in Concept storage are marked as irrelevance concepts. User interest is move to focus on "*Tsurugi apartment*" concept (as show in Figure 3).

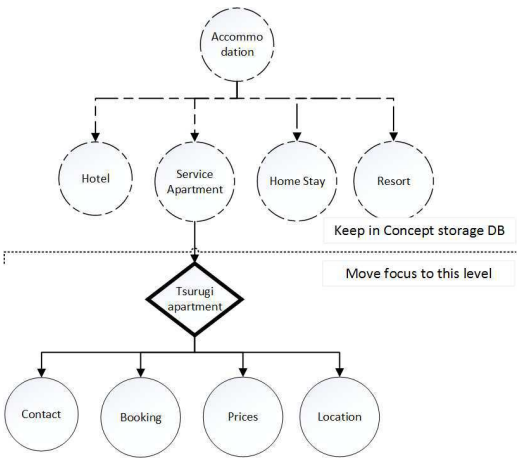


Figure 3. Example of feedback instance and related concepts

According to a propose system, feedback (link) of triples will carry a relate concepts to next loop. By starting at Semantic Query Expansion modules. In contrast to traditional approaches, all relevance triples will be formulated without a list of unused (irrelevance) concepts (such as Hotel, Home Stay, and Resort). The next possible triples are show as follow:

(Tsurugi Apartment, has, contact number)
(Tsurugi Apartment, has, price)
(Tsurugi Apartment, has, location)
(Tsurugi Apartment, has, booking process)
…

In this example, the final interest of searcher may be a Tusrugi apartment contact number.

Information interest is subjective, user need to identify interested information by their experience. As shown in this example, "Tsurugi Apartment contact number" is different from "Hotel near JAIST". Meanwhile, users do not know exactly what information that they looking for, until the exactly information arise. It seem impossible that user can state what they looking for at the first time, they need an iterative process to clarify what they interest.

5 Discussion and future work

In this paper, we propose the concept model of Semantic Query Expansion framework. By use concept driven (rely on specific Domain ontology) and user feedback. Combine with a nature of iterative search process. In traditional approaches, only generic model of query expansion is handled. However, personal interest is subjective. Even same query from different user (or different search session) may intent to locate difference information. By the proposed framework, we can achieve a tailor made query expansion according to user live interest. This framework will implement and practices to examine the usability in the future work.

Acknowledgement

This research was conducted under a grant in the SIIT-JAIST-NECTEC Dual Doctoral Degree Program.

References

[1] J. Wu, I. Ilyas, and G. Weddell, A study of ontology-based query expansion, *Technical report CS-2011-04*, University of Waterloo, 2011.

[2] A. Spink, D. Wolfram, M. B. Jansen, and T. Saracevic, Searching the web: The public and their queries, *Journal of the American society for information science and technology*, 52(3):226–234, 2001.

[3] S. Dumais, E. Cutrell, J. J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins, Stuff I've seen: a system for personal information retrieval and re-use, *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 72–79, 2003.

[4] M. Henzinger, Search Technologies for the internet, *Science*, 317(5837): 468–471, Jul. 2007.

[5] B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka, In Google we trust: users' decisions on rank, position, and relevance, *Journal of Computer-Mediated Communication*,12(3):801–823, Apr. 2007.

[6] G, Salton, and M, Mcgill, *Introduction to Modern Information Retrieval*. McGraw Hill, New York, NY, 1983.

[7] M. Mitra, A. Singhal, and C. Buckley, Improving automatic query expansion, Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, pages 206–214, 1998.

[8] J. Wei, S. Bressan, and B.C. Ooi, Mining term association rules for automatic global query expansion: methodology and preliminary results, *Proceedings of the First International Conference on Web Information Systems Engineering*, pages 366–373, Hong Kong, China, 2000

[9] E. M. Voorhees, Query expansion using lexical-semantic relations. *Proceedings of ACM SIGIR*, Dublin, Ireland, pages 61–69, 1994.

[10] S, Robertson, On relevance weight estimation and query expansion. *Journal of Documentation*, 42(3):182-188, 1986.

[11] L. Fitzpatrick, and M. Dent. Automatic feedback using past queries: Social

searching?, Proceedings of the 20th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, pages 306-313, 1997.

[12] C. Orasan, V. Pekar, and L. Hasler: A comparison of summarisation methods based on term specificity estimation, *Proceedings of LREC 2004*, Lisbon, Portugal, pages 1037–1041, 2007.

[13] B. Billerbeck, F. Scholer, H. E. Williams, and J. Zobel, Query expansion using associated queries, *Proceedings of the twelfth international conference on Information and knowledge management*, pages 2–9, 2003.

[14] H. Joho, M. Sanderson, and M. Beaulieu, A study of user interaction with a concept-based interactive query expansion support tool, *Advances in Information Retrieval*, Springer, pages 42–56, 2004.

[15] K. Järvelin, J. Kekäläinen, and T. Niemi, ExpansionTool: Concept-based query expansion and construction, *Information Retrieval*, 4(3–4):231–255, 2001.

[16] G. Leroy, K. M. Tolle, and H. Chen, Customizable and ontology-enhanced medical information retrieval interfaces, 1999.

[17] E. Mena, V. Kashyap, A. Illarramendi, and A. P. Sheth, Imprecise answers in distributed environments: Estimation of information loss for multi-ontology based query processing. *International Journal of Cooperative Information Systems*, 9(4):403-425, 2000.

[18] R. Chauhan, R. Goudar, R. Rathore, P. Singh, and S. Rao, Ontology Based Automatic Query Expansion for Semantic Information Retrieval in Sports Domain, *Eco-friendly Computing and Communication Systems*, Springer, pages 422–433, 2012.

[19] J. Bhogal, A. Macfarlane, and P. Smith, A review of ontology based query expansion, *Information Processing & Management*, 43(4):866–886, Jul. 2007.

[20] WordNet. Available online: http://wordnet.princeton.edu, accessed on 13 March 2010.

Color-based Ranking Method for Web documents

Yurika Moriwaki† **Hiroshi Honda‡** **Kosuke Takano†**

†Department of Information and Computer Sciences
Kanagawa Institute of Technology, Japan
‡Graduate School of Information and Computer Sciences
Kanagawa Institute of Technology, Japan
{s1021152, s1385018}@cce.kanagawa-it.ac.jp, takano@ic.kanagawa-it.ac.jp

Abstract

In this paper, we present a color-based ranking method for web documents. In our approach, we leverage color information of a web document, not only textual information, for ranking Web documents. By using user's preference color, which is automatically profiled into a database, our ranking method allows a user to obtain Web documents according to the user's color taste. From the experimental results, we confirm the feasibility of our color-based ranking system.

Keywords: Web, Color, Information retrieval, Ranking, Preference

1 Introduction

Presently, many web search engines ranks web documents based on textual information of web documents, where a user inputs query terms into a search engine, and receives the ranking results through a web browser interface based on the similarity score.

However, we think that it is significant to consider the readability of web documents as the rank score, not only textual information, for ranking web documents. For example, when a user receives news articles from information recommendation systems [3], it would satisfy the user if the system provides news articles with the good readability [4] in terms of the color combination, contrast, and brightness. In addition, it is important for users who have particular color vision such as color blindness to obtain web documents with their acceptable colors from a search engine. For example, if a user has color blindness for red color, the web documents that uses red color in large part should be excluded from the result, and instead suggest the web documents using another colors that the user can accept .

In this study, we assume that some users have a color image associated with the content of a Web site. For example, if a female user likes clothes in Spring-associated color such as yellow and pink, she would often browse Web sites of apparel products where clothes in yellow and pink are displayed. In the browsing process, she might cherish the Spring-color image for the Web sites of online apparel shops. In this point, if she has registered "fashion" as a keyword for a recommender system, the system should preferentially suggest the Web sites that display apparel products in the Spring-color image such as yellow and pink in the recommendation.

In this paper, we present a color-based ranking method for web documents. In our approach, we leverage color information of a web document, as well as textual information for ranking Web documents. By using user's preference color, which is automatically profiled into a database, our ranking method allows a user to obtain Web documents according to the user's color taste. From the experimental results, we confirm the feasibility of our color-based ranking system.

2 Color-based Ranking

Figure 1 shows overview of our system. Firstly, a user's color profile is constructed from the color information of web pages that the user is browsing through a web browser. Meanwhile, a web crawler with a color extractor indexes web documents using color metadata as well as term metadata. Using the user's color profile and the color metadata attached to each web document, our color-based search engine ranks web documents according to user's color taste.

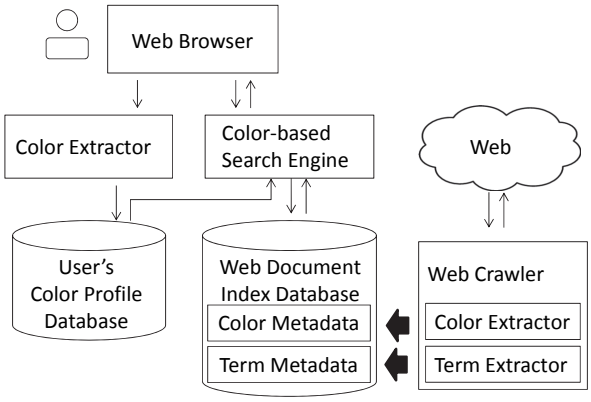


Figure 1. Overview of System

2.1 Construction User's Color Profile

Through a continuous user's browsing of web documents, the color information c_w of the web documents that the user is viewing is accumulated into a user color profile. The user's color profile is represented as a color-based feature vector $\mathbf{c}_u$, where n-colors definition (c_1, c_2, …, c_n) is used as the schema of feature vector, as follows:

$$\mathbf{c}_u = \lfloor c_{u1}, c_{u2}, \cdots, c_{un} \rfloor \tag{1}$$

where c_{ui} is an accumulated pixel number of the corresponding color c_i.

In the current implementation of our system, the web document that a user is viewing is captured as an image format, then the color information c_w of the web page is extracted from the captured image. Therefore, the color information is extracted not only for text, but also from images in the web pages that the user has browsed.

2.2 Extraction of Color Metadata

Color metadata is extracted from each web document by a color extractor in a web crawler. Using the same color scheme of $\mathbf{c}_u$, color metadata $\mathbf{c}_x$ of web document d_x is represented as a feature vector. All colors in d_x are reduced into n-colors schema by assigning a color to the similar color c_i.

$$\mathbf{c}_x = \lfloor c_{x1}, c_{x2}, \cdots, c_{xn} \rfloor \tag{2}$$

where c_{xi} is a pixel number of the color c_i used in d_x.

2.3 Extraction of Term Metadata

In order to extract term metadata from web documents, our method extracts term sets T_x from each web document d_x. T_x includes the m terms appearing in d_x.

$$T_x = \{ t_{x1},\ t_{x2},\ t_{x3},\ \cdots,\ t_{xm} \} \tag{3}$$

where t_{xi} (i = 1, 2, …, m) is a term included in T_x. The total term set T of all terms appearing in each T_x is represented as follows:

$$T = \bigcup_{i=1}^{x} T_i \tag{4}$$

By using each term included in the term set T and each web document d_x, our method creates a document vector space [1] and a vectorized d_x.

$$\mathbf{d}_x = \lfloor u_{x1}, u_{x2}, \cdots, u_{x|T|} \rfloor \tag{5}$$

where u_{xi} is a term frequency of t_i appeared in a web document d_x.

2.4 Ranking of Web Documents

In the ranking process, a user's query consists of a color-based feature vector $\mathbf{c}_u$ that is obtained from the user's color profile and a query term vector $\mathbf{q}$ of the user's input query terms q.

In our ranking method, firstly, the color-based similarity score S_c between c_u and c_x is calculated using the cosine measure [1] as follows:

$$Sc(c_u, c_x) = \frac{\mathbf{c}_u \cdot \mathbf{c}_x}{|\mathbf{c}_u||\mathbf{c}_x|} \tag{6}$$

Then, the term-based similarity score S_t between user's query terms q and d_x is calculated as follows:

$$St(q, d_x) = \frac{\mathbf{q} \cdot \mathbf{d}_x}{|\mathbf{q}||\mathbf{d}_x|} \tag{7}$$

Using both S_c and S_t, the ranking score S for each web document d_x is calculated.

$$S = \alpha \cdot S_c + (1 - \alpha) \cdot S_t \tag{8}$$

Where α $(0 \le \alpha \le 1)$ is a parameter for balancing S_c and S_t.

Finally, our method ranks each web document based on the score S so that the user can obtain web documents according to his/her color taste.

3 Preliminary Experiments

In the preliminary experiment, we confirm the feasibility of our color-based ranking method for web documents.

Table 1. Example of Metadata

Doc ID	Genre	Example of Metadata
d_{1a} d_{1b} d_{1c}	Fashion	coordination, fresh, girl, ecology, knit, dungaree, beige, navy
d_{2a} d_{2b} d_{2c}	Cook	cook, recipe, ingredient, butter, pasta, penne, gnocchi, rice, bean
d_{3a} d_{3b} d_{3c}	IT business	solution, agile, software, development, network, service
d_{4a} d_{4b} d_{4c}	Computer Lab	research, publication, lecture, member, database, mobile
d_{5a} d_{5b} d_{5c}	Music	lesson, service, support, audio, chorus, instrument, mobile, smartphone

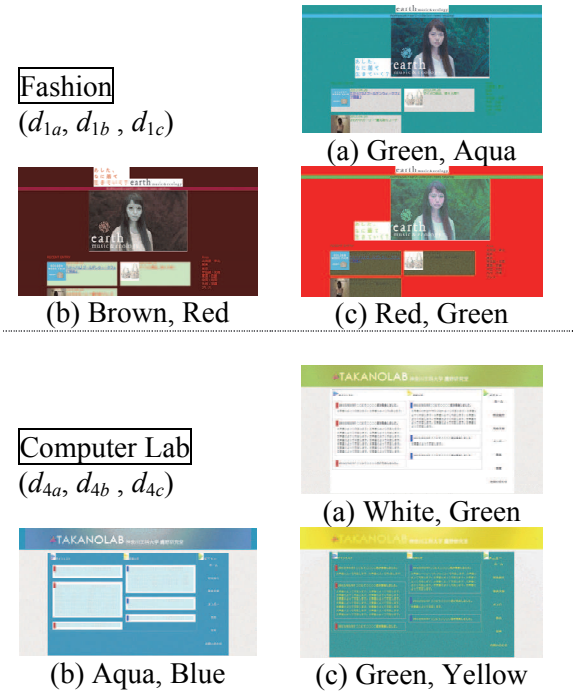


Figure 2. Example of Colors used in Web Documents

In our prototype, we implemented a color profiler for web documents browsed by a user using the Google Chrome extension. The color profiler works in background, and automatically extracts the color information from the captured image of the web documents browsed by a user. The extracted color information of each browsed web documents is accumulated into the user color profile as described in Section 2.1.

In addition, we prepared 5 web documents d_1-d_5, where each web document has three different color sets. Therefore, we have totally 15 web documents. Table 1 and Figure 2 show examples of metadata and web documents with different color sets, respectively.

In the experiment, we evaluate the ranking capability of our method. We used the following query term set, and user color profiles 1 and 2.

Query: mobile, system

User color profile 1: (Main color) Red, khaki
User color profile 2: (Main color) Green, white

Tables 2 and 3 show the ranking result for the user color profiles 1 and 2, respectively, for the same query terms "mobile, system". In the results, bigger α is, stronger the color impact is for calculating the ranking score. For example, α=0.0 means that only the term-based similarity is used for the ranking, and α=0.5 means that both the term-based similarity and color-based similarity is equally used for the ranking.

From Tables 2 and 3, we can confirm that as α is getting bigger, web documents with user's color taste are getting searched in the top rank. For example, in Table 2, d_{5c} (main color: red, aqua) is ranked up 1st (α=0.3) from 9th (α=0.0) for user's color taste: red, khaki. In addition, from metadata attached to d_{5c}, d_{5c} is related to a topic of music including mobile devices such as a smartphone. Therefore, this result shows the feasibility of our approach that our method allows a user to obtain web documents in terms of both textual information and color information of the web documents.

Table 2. Ranking Result (1): User color profile 1

Rank	α			
	0.0	0.3	0.5	1.0
1	d_{4a} 0.182	d_{5c} 0.278	d_{1c} 0.433	d_{1c} 0.867
2	d_{4b} 0.182	d_{1c} 0.260	d_{5c} 0.353	d_{5c} 0.540
3	d_{4c} 0.182	d_{3b} 0.239	d_{3b} 0.280	d_{2b} 0.445
4	d_{3a} 0.176	d_{4a} 0.217	d_{4a} 0.240	d_{3b} 0.385
5	d_{3b} 0.176	d_{4b} 0.193	d_{2b} 0.222	d_{2a} 0.301
6	d_{3c} 0.176	d_{3a} 0.174	d_{4b} 0.200	d_{4a} 0.298
7	d_{5a} 0.166	d_{4c} 0.166	d_{3a} 0.173	d_{2c} 0.246
8	d_{5b} 0.166	d_{5a} 0.155	d_{4c} 0.155	d_{4b} 0.219
9	d_{5c} 0.166	d_{5b} 0.147	d_{2a} 0.150	d_{3a} 0.169
10	d_{1a} 0.000	d_{3b} 0.140	d_{5a} 0.148	d_{5a} 0.130
11	d_{1b} 0.000	d_{2b} 0.133	d_{5b} 0.134	d_{4c} 0.129
12	d_{1c} 0.000	d_{2a} 0.090	d_{2c} 0.123	d_{5b} 0.101
13	d_{2a} 0.000	d_{2c} 0.074	d_{3c} 0.117	d_{1a} 0.075
14	d_{2b} 0.000	d_{1a} 0.022	d_{1a} 0.037	d_{3c} 0.057
15	d_{2c} 0.000	d_{1b} 0.010	d_{1b} 0.017	d_{1b} 0.035

Table 3. Ranking Result (2): User color profile 2

Rank	α			
	0.0	0.3	0.5	0.7
1	d_{4a} 0.182	d_{1a} 0.292	d_{1a} 0.486	d_{1a} 0.973
2	d_{4b} 0.182	d_{4c} 0.279	d_{4c} 0.344	d_{4c} 0.507
3	d_{4c} 0.182	d_{4b} 0.232	d_{4b} 0.265	d_{4b} 0.349
4	d_{3a} 0.176	d_{4a} 0.185	d_{4a} 0.188	d_{2a} 0.194
5	d_{3b} 0.176	d_{3a} 0.158	d_{3a} 0.145	d_{4a} 0.193
6	d_{3c} 0.176	d_{5c} 0.147	d_{5c} 0.134	d_{2b} 0.177
7	d_{5a} 0.166	d_{5a} 0.141	d_{5a} 0.124	d_{2c} 0.150
8	d_{5b} 0.166	d_{3b} 0.137	d_{5b} 0.116	d_{3a} 0.114
9	d_{5c} 0.166	d_{5b} 0.136	d_{3b} 0.111	d_{5c} 0.102
10	d_{1a} 0.000	d_{3c} 0.134	d_{3c} 0.107	d_{5a} 0.082
11	d_{1b} 0.000	d_{2a} 0.058	d_{2a} 0.097	d_{5b} 0.065
12	d_{1c} 0.000	d_{2b} 0.053	d_{2b} 0.088	d_{3b} 0.045
13	d_{2a} 0.000	d_{2c} 0.045	d_{2c} 0.075	d_{3c} 0.037
14	d_{2b} 0.000	d_{1c} 0.004	d_{1c} 0.007	d_{1c} 0.015
15	d_{2c} 0.000	d_{1b} 0.004	d_{1b} 0.006	d_{1b} 0.013

4 Conclusion

In this paper, we have presented a color-based ranking method for web documents. From the experimental results, we confirm the feasibility of our color-based ranking system.

Our system mainly consists of a user's color profile database, a color-based search engine, and a web crawler with color-metadata extractor from web documents. Using the user's color profile and the color metadata attached to each web document, our color-based search engine ranks web documents according to user's color tastes, not only text-based query terms.

Presently, we are implementing the framework of our system including a web crawler with a color indexer as shown Figure 1, and we will show the feasibility and the effectiveness of the proposed system in our future work.

References

[1] Baeza-Yates, R. and Ribeiro-Neto, B. *Modern Information Retrieval*, Addison Wesley, 1999.

[2] Deerwester, S., Dumais, S., Furnas, G.W., Landauer, T.K. and Harshman, R., *Indexing by latent semantic analysis*, Journal of the American Society for Information Science, Vol. 41, No. 6, pp.391-407, 1990.

[3] Jannach, D., Zanker, M., Felfernig, A., Friedrich, G, *Recommender Systems: An Introduction*, Cambridge University Press, 2010.

[4] Hearst, M. A., *Search User Interfaces*, Cambridge University Press, 2009.

ACIS 2013

Asian Conference on Information Systems 2013

Special Session on IT Management

Assessing E-Business Deployment Framework for Thai SMEs:

—— The Case of Thai Restaurants in Melbourne, Australia ——

Suttisak Jantavongso **Kulthon Kasemsan**
Faculty of Information Technology
Rangsit University
suttisak.j@rsu.ac.th

Abstract

This paper is one of series of working papers on the 'E-Business Deployment Framework for Thai SMEs' aka (EBDF) by the authors. Thai Small to Medium Enterprises (SMEs) play an extremely important role in the Thai economy. Over the last decade, many SMEs have added Electronic Business (e-business) to their traditional business model in an attempt to increase their competitiveness and chances of survival. In the case of Thai SMEs, e-business offers the potential of advancing directly to a world competitive basis without having to follow a traditional development path involving considerable time and expense. A notable characteristic of existing Western e-business models is their concentration on the technology of designing and implementing such models. However, a far more complex situation arises when considering the adoption of e-business by SMEs in Thailand. In response to the many potentially complex issues which arise, this study has attempted to access the applicable of the EBDF in the context of Thai restaurants in Melbourne. The EBDF was designed to address many of the issues which arise in the adoption of e-business by Thai SMEs. Unlike many existing Western e-business models, the EBDF provides an holistic model which not only addresses the technology but also addresses a range of non-technical issues associated with the adoption of e-business. Included in the range of non-technical issues considered are elements such as: complexity, trust, confidentiality, human resources, supporting infrastructure, management and finance. Six Thai restaurants classified under the SMEs category in Melbourne were participated to assess the applicable of the EBDF. This paper presents a discussion on the findings. Mixed methods approach involving observations, interviews, and document reviews to gather relevant data were employed. The observations took place in February 2013. The study also investigated the possibility of replicated, and comparative results to provide extra clarity and insight. The preliminary findings were aligned with the authors' existing studies, and that the EBDF with some contextual modification was applicable for the Thai SMEs operating overseas.

Keywords: E-Business, Deployment Framework, ASP, SME, PSP, SME Enterprise Software

1. Introduction

The role of Small to Medium Enterprises (SMEs) is crucial to the economic development and performance of any country. Literature findings by the authors reveal that SMEs around the world are turning to Electronic Business (e-business) technology for solutions to increase their competitiveness and chances of survival. Many SMEs have already added electronic forms of business to their traditional business model [1].

For Thai SMEs, the adoption of e-business offers the potential of advancing directly to their customers. The study by Blashki & Jantavongso [2] cited in Regan & Wymer [3] also identified the desire to improve marketing opportunities by Thai SMEs as the number one ranked reason for adopting e-business. While e-business technologies offer many and diverse opportunities to strengthen their businesses, many issues and constraints need to be overcome before such opportunities can be fully exploited. The 'E-Business Deployment Framework for Thai SMEs' aka (EBDF) which was developed by the authors has addressed these limitations.

In addition, the Thai government is keenly aware of this, and has adopted a number of proactive policies to enhance their development. Thai SMEs are also keen to adopt e-business technologies to remain competitive in this digital economy. However, lack of both financial and human resources is a major impediment.

The aim of the EBDF is to facilitate e-business implementation among SMEs through a holistic approach which addressed not only the technological aspects, but also the managerial issues within the specific social and cultural context. The EBDF specifically developed for Thai SMEs, to enable them to overcome the lack of resources. The EBDF allows Thai SMEs to distribute the costs of e-business by sharing resources. It has five major components: (1) Application Service Provider (ASP), (2) Professional Service Provider (PSP), (3) SME Enterprise Software, (4) Government Involvement, and (5) Trust Facilitation. The complete details of the EBDF was published by the authors in the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA) 2006 [4]

Since 2002, the EBDF has been developed, revised, evaluated, and validated a through a sound scientific process of a theory building and testing. The EBDF also has statistically significant level of acceptance able to support a

wide range of Information System (IS) research, especially through the research procedure in which a better understanding of what contributes to the success of e-business among SMEs in Thailand has been demonstrated through the authors' publications [2, 4, 5, 6, 7, 8, 9].

In 2007, the authors had made a number of suggestions for the future research of the EBDF [1]. While the framework is found to be significant, and contributed to the growing body of research into issues of relevance to e-business and SMEs. As in all studies, the EBDF does have some issues. One of these issues was that the framework does not have complete acceptance (100 per cent); and there was room for consideration. Moreover, the framework was targeted only the SMEs in Thailand. An interesting area of future research was to test the adoption level of the framework in other cultural and economic settings. The remaining of this paper presents the results of the findings.

2. Literature Review

As indicated, Thai SMEs are of interest of this study; and of particular interest region is Melbourne which provides a specific point of reference. Within Melbourne, Thai restaurants are closely associated as Thai SMEs. Following Jantavongso [10], Thai cuisines are very popular throughout Australia, due to the variety of tastes. Thai cuisines are foods of health, consisting of herbs as main ingredients. For that reason, numbers of Thai restaurants have been expanded around Australia, and having tendency for the continued increasing. Thailand and Australia have also established the relation over a long period of time. These days, there are more than 1,000 Thai restaurants in many capital cities in Australia, and 300 Thai restaurants in Melbourne alone.

Accordingly, all of the participated restaurants in this study are belong to the SMEs category. The classification was followed the recommendation by Othman & Don [11] in their studies of the business performance in a restaurant industry.

2.1 Thai SMEs in Melbourne

Following Jantavongso [10], majority of Thai SMEs in Melbourne are Thai restaurants. One of the longest established Thai Restaurants in Melbourne was opened since 1983. Initially, Thai restaurant owners, managers, chefs, cooks, and staff first came to Melbourne to study. Nearly all of the people who work in the Thai restaurants are Thai descents. Working at a Thai restaurant had provided extra income while studying in Melbourne. Most of the Thai restaurants owners had worked between five to ten years in the restaurants to develop their skills from general staff, cook, chef, manager, and then owner respectively.

Furthermore, the locations of the Thai restaurants in Melbourne are scattered around twenty-seven (27) kilometers radius from the Central Business District (CBD). The Thai restaurants in Melbourne are varied in size from large to small, depending on the number of seats, and their management capabilities. Majority of the Thai restaurants in Melbourne are operated as small family businesses. The owners play an important decision making role in running the businesses including the adoption of e-business.

Melbourne was selected as the city under study due to a number of reasons. One of these reasons was the level of 'digital economy ranking' formally known as 'e-readiness' of Australia. Thus, there have been a declared interest within the Melbournian communities in promoting and supporting Thai SMEs. Thais, Thai restaurants, and City of Melbourne play an important role in promoting Thai culture, and authentic Thai cuisine to the Melbournians. For example, the 'Thai Culture and Food Festivals' are held annually since 2004 at the 'Federation Square'. The festival attracted more than 64,000 visitors in 2012, making this one of the top ten Melbourne events of the year [12].

In addition, the 'Thai Kitchen World Cuisine' is another initiative by the Thai government to support Thai SMEs operating overseas. The Thai government also launched the 'AROY @ Melbourne' campaign in 2005 at the festival to promote the consumption of Thai food. These initiatives help Thai SMEs to build a network resources to overcome their major obstacles and create new market opportunities. As mentioned, the desire to improve marketing opportunities by Thai SMEs is the number one ranked reason for adopting e-business. As such, the Thai government must play an important role; and active Thai government support is essential if e-business is to be adopted by Thai SMEs operating overseas and succeed.

2.2 Thailand's digital economy ranking

A notable characteristic of existing e-business models has been an overt concentration on the value of technology. However, many more issues arise within a Thai context. Lack of resources remains a major impediment to Thai SMEs' deployment of e-business and establishment of a presence online [1].

Moreover, Thailand cannot be considered to have a high level of 'digital economy ranking', given the developing nature of the Thai economy and the labour intensive semirural nature of Thai SMEs. Despite this assertion, Thai SMEs are readily adopting e-business while not having a high level of 'digital economy ranking'. The implication here is that the EBDF must harness this enthusiasm for e-business by Thai SMEs, while not falling into the trap of overselling its potential. Moreover, because of the very size and nature of Thai SMEs, they are unlikely to possess the resources needed to adequately address the issues and factors, and this situation poses a substantial risk of failure.

Accordingly, 'digital economy ranking' measures the capability of a country's business environment to seize Internet-based commercial opportunities. A large number of countries that have a high level of 'digital economy ranking' are Western countries including Australia. Nonetheless, a high level of 'digital economy ranking' is

not exclusively a Western attribute and Asian countries, such as Singapore, also have a high level [13].

3. E-Business Deployment Framework for Thai SMEs (EBDF)

In response to the many potentially complex issues which arise, the authors proposed an 'E-Business Deployment Framework' (EDBF) in 2007. The EBDF was designed to address many of the issues which arise in the adoption of e-business by Thai SMEs. The EBDF is developed through both (a) an intensive literature review, to ensure that they are grounded within the theoretical perspective of the critical e-business success factors; and (b) in depth face-to-face interviews with the total of 160 (one hundred and sixty) Thai SMEs, to ensure its relevance to the context of Thai SMEs. The EBDF was shown to have a real value, and it had a statistically significant level of acceptance of 77 per cent, which did not vary significantly between different business sectors and within various regions of Thailand.

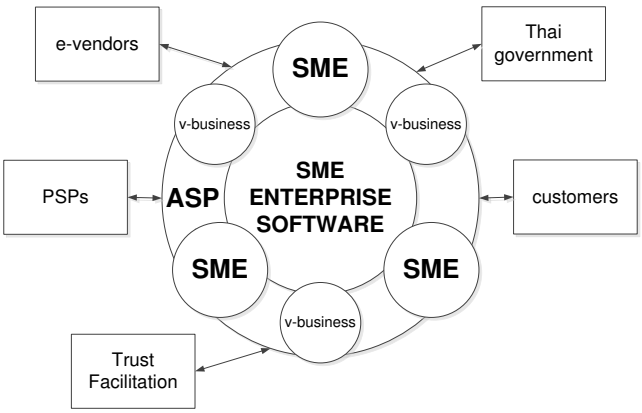


Figure. 1. Key elements in the EBDF

The framework integrates external service providers, ASPs, SME enterprise software, government involvement, and trust facilitation. The heart of the framework is the 'SME Enterprise Software' module. This software module is operation-centric which is owned, hosted and operated by the ASP to assist a group of SMEs to conduct their day-to-day business operations. The software module also enables back-office integration. Figure 1 demonstrates the inter-relationship between ASP and SMEs. SMEs share a common team of PSPs through the ASP. This team of providers includes management consultants, banking, online purchase transaction service providers, training experts, accountants, lawyers, IT professionals, and benchmarking specialists. However, the benefits of this arrangement are not necessarily limited to accessing resources. As noted, the framework involves cooperative arrangements. SMEs also benefit from the combined purchasing power of all members serviced by the ASP and ASPs are able to negotiate better offers and prices with electronic vendors (e-vendors). Another important stakeholder in the framework is the virtual businesses (v-businesses). According to Barnes [14] , a v-business operates its entire business on the Internet and usually operates without a conventional physical commercial office presence known to its electronic customers (e-customers).

4. Research Methodology

4.1 Research Design

The study has adopted the framework proposed by Creswell [15] which conceptualized Crotty's [16] framework. Following Creswell [15]'s framework, the mixed methods approach which based knowledge claims on pragmatic grounds was employed. The data collection involved gathering both numeric information as well as text information so that the final findings represent both quantitative and qualitative information.

Due to the nature of this sutdy, a survey approach and field research had been adopted [17]. Survey research was suitable for identifying the frequency of certain characteristics of Thai SMEs in Melbourne. This allowed the authors to relate various characteristics in order to explain a phenomenon. Accordingly, the field research allowed the authors to gain firsthand knowledge. The aim was to gather information without influencing the environment. The difficulty faced by the authors was determining when and what observations to record. Also, an opportunity had to be available and accessible to conduct the observations.

4.2 Sampling Framework

Following Jantavongso [10], a 'purposive sampling' method is employed in this study. Six Thai restaurants were selected, interviewed, and thoroughly observed. This paper presents preliminary findings of one of these restaurants as a pilot case. The observations took place in February 2013. Like Chanvarasuth [18], the size of SMEs was emphasized, and used to determine the sampling framework. The size of restaurant was determined by its number of seats. A 71 seats and above restaurant is classified as a large-sized, a 31-70 seats is a medium-sized, and a 30 seats or less is a small-sized restaurant respectively. The sample of the six Thai SMEs was divided equally. Participation in this research was entirely voluntary.

4.3 Description of the Subject

'Blue River Thai Restaurant' or BRT, is located at 239 Canterbury Road, Canterbury, Melbourne, Australia. BRT is established in 2009 and has 68 seating. Currently, BRT offers takeaway, home delivery, functions, events, and dine in services to its customers. For takeaway and home delivery services, BRT offers both telephone and online ordering options. BRT offers its online ordering since 2011.

4.4 The Hypothesis

Although the EBDF is targeted the at SMEs in Thailand, the framework is expected to be applicable to Thai SMEs operating in other countries, perhaps with some contextual modification, or at least serve as the catalyst for a similar type of study. Therefore, it is envisaged that, the EBDF is also able to facilitate Thai SMEs in Melbourne to successfully implement e-business.

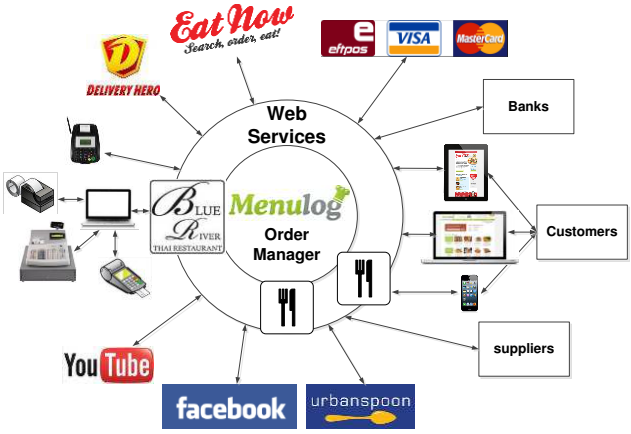


Figure. 2 Mapping of Blue River Thai Restaurant and the E-Business Deployment Framework

5. Findings

Even though, the literature review by the authors suggested that Thai SMEs and Thai SMEs in Australia are similar in aspects of success criteria used to gauge business performance, and characteristics used to differentiate them from large enterprises; differences do exist. A number of key findings emerged, for example, the underlying business culture in Thailand is a system of 'patronage' based on adherence to personal relationships rather than to regulation by an institution. SMEs in Thailand are labor intensive. The extensive role of the Thai government is required to provide Thai SMEs with assistance. This is not a case for Thai SMEs in Australia.

As indicated, lack of both financial and human resources is a major impediment for both Thai SMEs and Thai SMEs operating overseas to embrace e-business. However, Australian customers feel more comfortable regarding security and privacy in e-business. Thus, it is apparent that most Australian organisations already made substantial investment in e-business resources [19].

5.1 ASP and Professional Service Provider

An ASP model emerged in the late 1990s, it offers benefits of providing professional services such as management consultants, training experts, accountants, lawyers, IT professionals, and benchmarking specialists required to conduct e-business [20]. This is in addition to the traditional sharing of hardware, software and technical support infrastructure services among SMEs, in particular when running an e-business. Thus, it is not uncommon for an ASP model to be included in an implementation or a deployment of e-business in SMEs.

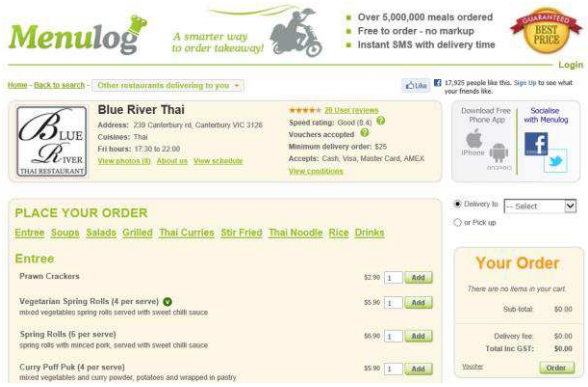


Figure. 3 Takeaway & home delivery websites hotes by Menulog

BRT acquires services from the ASPs and Professional Service Providers (PSPs) via three main online takeaway and home delivery service providers: (1) 'Menulog Pty Ltd', (2) 'Delivery Hero Australia', and (3) 'Eat Now Services Pty Ltd'. These companies provide BRT with an e-business platform focused on the takeaway food sector. The e-business platform is centred around a website and mobile app that enable customers to browse menus and order takeaway food from BRT as well as other restaurants in their area (see Figure 2). Customers can also filter results by cuisine or payment type.

For example, 'Menulog Pty Ltd' provides BRT with takeaway and home delivery websites (see Figure 3). This allows BRT to obtain increase orders from new customers. Furthermore, there is no monthly fees, no joining fees, and no contract period. It is extimated average increase in ordering for BRT of 20% since 2011.

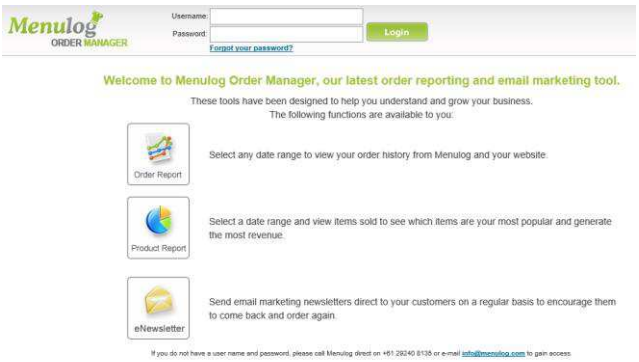


Figure. 4. Menulog order manager system

5.2 SME Enterprise Software

One of the important issues to be addressed was the need for Thai SMEs to build integrated business processes within and across organizations. It is essential that each SME's websites be supported with a mature back office operation and that the e-business solutions be seamlessly integrated. The framework was designed to make use of an 'operation-centric' e-business system to be provided through an ASP.

'Menulog Pty Ltd' provides BRT with online the Menulog Order Manager (MOM) system (see, Fugure 4) at no cost. It has found it to be a very useful addition to BRT. It has eliminated the need for fax and any follow-up telephone calls. The system is very easy to use. Orders are clearly displayed on the screen complete with the customer's details. Staff is able to instantly confirm orders with a simple touch of the screen. Staff is also able to view order history. The system allows compilation database of customers to promote BRT.

5.3 Government Involvement and Trust Facilitation

As indicated, trust occupies a particularly important place in Thai business culture, more importantly when conducting a business transaction online. Accordingly, the Thai government is required to play a major role in establishing and nurturing e-business under the framework. The Thai government needs to provide ethical e-business standards governing the behaviour of each party involved and to take an active role in promoting, educating and assuring the effectiveness of e-business to the public.

Furthermore, the EBDF emphasizes the complex issue of trust must be considered in the context of the Thai patronage system, and it is one of much greater significance in Thailand than would be expected in Australia. The framework integrates the technological, the value proposition, and the trust basis within a single coherent framework

Nonetheless, the government involvement and trust elements become less significant in the context of Thai SMEs in Melbourne. Accordingly, there is a already high level of awareness of Australian consumer protection laws and most consumers are confident that businesses do the right thing. Consumers are aware of consumer regulators, and about half of consumers will contact a regulator about a consumer problem. There is broad awareness of dispute resolution services for consumer problems and most consumers and businesses would use a dispute resolution service if they needed to [21].

5.4 Web 2.0 and Social Networking

'Facebook' is a Web 2.0 and one of the world's most popular social networking websites [22]. It provides platform for sharing information, videos, photos, and discussions. With approximately 600 million members, Facebook has provided the best opportunity for BRT. Apart from being an effective marketing tool, Facebook is a new trend of innovation for restaurant businesses.

Due to the resource limitations of BRT, traditional modes of marketing like television advertisings (ads), print media and outdoor media alone cannot be relied upon to ensure it reach its target customers. Facebook has two primary types of ads: standard ads, and socials ads [23]. Standard ad is a normal type of click through ad on the Internet. Facebook social ads allow BRT to become part of people's daily conversations (see Figure 5).

Figure. 5 Existing BRT's facebook campaign

Moreover, there are two possible options for BRT to create an ad on Facebook: (a) cost per click (CPC) or (b) cost per thousands (CPM). The minimum of CPC is $0.01. Alternatively, CPM ads cost $0.15 for the ad space and $0.15 for news feeds. In addition, the minimum daily budget for both CPC and CPM is $5 per day. BRT would need to specify the maximum it wants to pay for each click or each one thousand impressions.

BRT is also paid attention on 'YouTube'. YouTube marketing can also become one of crucial promotion channels for BRT. YouTube can also act as a distribution channel for BRT; and provides a forum for its members and potential members to connect, inform, and inspire others across the Internet. YouTube can be setup, create and customise a BRT brand channel [24]. It is also be able to distribute videos to other sites. In summary, BRT does not ignore the potential benefits of YouTube and Video Advertising to reach its target customers. Thus, it is free to place a promotion video on YouTube.

6. Work in Progress

The authors have been engaged in a number of studies to investigate the applicable of the EBDF among Thai SMEs operating in other countries. While, this study was targeted at the six Thai SMEs in Melbourne, Australia; the paper has reported only the preliminary finding base on one participant. The authors' primary objective is to understand how the framework would fit into the Thai restaurants, and determine what the factors are that hinder them in implementing e-business. An interesting area of future research would be to test the acceptance level of the framework in the Australian cultural and economic settings. The detailed findings will be published in future publications.

7. Conclusion

In this paper, the authors had briefly reintroduced the EBDF. The framework was designed based on specific objectives and strategies to ensure its relevancy to the context of Thai SMEs. The framework includes the employment of ASP and PSPs to encourage sharing of resources; the use of SME enterprise software; pro-active role of the Australian government; and trust facilitation among all parties. The authors employed a mixed methods involving observations, interviews, and document reviews. The observations took place in February 2013. While the framework was primary targeted the SMEs in Thailand; the findings indicated that the EBDF was applicable to Thai SMEs operating overseas with some contextual modification. Thai SMEs in Australia are also able to implement the EBDF as a feedback mechanism to investigate their own e-readiness; it can help SMEs to identify their strengths and weaknesses before proceeding with or evaluating their current e-business activity.

References

[1] S. Jantavongso, *Facilitating e-business implementation among Thai SMEs*, 2007, Monash University. p. xv, 257 leaves.

[2] K. Blashki and S. Jantavongso. *E-Business in SMEs of Thailand: a descriptive survey*. in *Proceedings of the 2006 Information Resources Management Association International Conference*. 2006. Washington DC: Idea Group Publishing.

[3] E. Regan and S. Wymer, *Influential factors in the adoption and use of e-business and e-commerce information technology (EEIT) by small & medium businesses.* Journal of Electronic Commerce in Organizations, 2011. **9**(1): p. 56-82.

[4] L.F. Sugianto and S. Jantavongso. *eBusiness Adoption Studies Focusing on Thai SMEs*. in *Proceedings of the 2006 International Conference on Computational Intelligence for Modelling, Control and Automation*. 2006.

[5] S. Jantavongso and K.Y.R. Li. *E-Business in Thailand: Social and Culture Issues*. in *Proceedings of the 2002 Information Resources Management Association International Conference*. 2002. Seattle.

[6] S. Jantavongso, R.K.Y. Li, and B. Tootell, *Culture and E-Business in Thailand*, in *Managing Globally with Information Technology*. 2003, IGI Global. p. 177-195.

[7] S. Jantavongso and R.K.-Y. Li. *A new age e-business model for SME*. in *Proceedings of the AusWeb03: the 9th Australian World Wide Web Conference*. 2003. Gold Coast.

[8] K.-H. Oh, S. Jantavongso, and R.K.-Y. Li. *A critique: factors contributed to the dotcom failures*. in *Proceedings of the 2004 Information Resources Management Association International Conference*. 2004. New Orleans.

[9] S. Jantavongso, L.F. Sugianto, and Y.R. Syaub. *E-business deployment framework for Thai SMEs*. in *Proceedings of the International joint Conference of e-CASE, e-Commerce, e-Administration, e-Society, and e-Education*. 2007. Hong Kong.

[10] C. Jantavongso, *A Case Study of the Business Management Development for Restaurants in Melbourne, Australia in the Context of Thai Cultural Dimensions*, 2012, Mahasarakham Universit. p. 225 leaves.

[11] A.R. Othman and M.S. Don. *Knowledge management and business performance of casual dining restaurants in Malaysia*. in *Proceedings of the Knowledge Management International Conference (KMICe) 2012*. 2012. Johor Bahru.

[12] R. Karis. *Thai culture and food festival 2013*. 2013 [cited 2013 March 14]; Available from: http://www.meldmagazine.com.au/2013/03/thai-culture-and-food-festival-2013/.

[13] Economist Intelligence Unit and IBM, *Digital economy rankings 2010: beyond e-readiness*, 2010, The Economist Group: London. p. 1-24.

[14] S. Barnes, *E-commerce and v-business: digital enterprise in the twenty-first century*. 2nd ed. 2007, Oxford: Butterworth-Heinemann. xxii, 380 p.

[15] J.W. Creswell, *Research design: qualitative, quantitative, and mixed methods approaches*. 3rd ed. 2009, Thousand Oaks: Sage Publications. xxix, 260 p.

[16] M. Crotty, *The foundations of social research: meaning and perspective in the research process*. 1998, St Leonards: Allen & Unwin. vii, 248 p.

[17] P. Kazanis, *Methodologies and tools for etransforming small to medium size enterprises*, 2004, University of Western Sydney. p. xi, 210 leaves.

[18] P. Chanvarasuth, *Adoption of E-Business by Thai SMEs*. World Academy of Science, Engineering and Technology 41, 2010. **65**: p. 741-746.

[19] M. Singh and J. Byrne. *E-business evaluation in Australia: a discussion of some preliminary findings*. in *Proceedings of the Fourth International Conference on Electronic Business*. 2004. Beijing.

[20] N.A.Y. Al-Qirim, *Electronic commerce in small to medium-sized enterprises: frameworks, issues and implications*. 2004, Hershey: Idea Group. xv, 370 p.

[21] Commonwealth of Australia, *Australian consumer survey 2011*, 2011, The Treasury: Parkes. p. xiv, 256 leaves.

[22] N. Carlson. *Goldman to clients: Facebook has 600 million users*. 2011 [cited 2013 15 March]; Available from: http://www.msnbc.msn.com/id/40929239/ns/technology_and_science-tech_and_gadgets/.

[23] S. Holzner, *Facebook marketing: leverage social media to grow your business*. 2009, Indianapolis: Que. xii, 273 p.

[24] G. Jarboe, *YouTube and video marketing: an hour a day*. 1st ed. 2009, Chichester: John Wiley. xxx, 474 p.

The actual situation about practical capabilities of information control in the certified nurse administrator education first and second level trainees

Takako Izumi† **Yukie Majima ‡**

†, ‡ Morinomiya University of Medical Sciences, Japan
‡Graduate School of Engineering, Osaka Prefecture University, Japan
{izumi}@morinomiya-u.ac.jp

Abstract

This study was conducted to reveal the present circumstances and problems related to practical capabilities of information control and clinical nursing among nurse administrators. Therefore, to elucidate the practical capabilities of information control and use among nurse administrators, we conducted a survey of "the practical capability of information control and use of nurse administrators" using self-administered questionnaire forms. Subjects were 123 people who took certified nurse administrator training hosted by Japanese Nursing Association during August 2011 – December 2011. Results revealed that most people considered that "capabilities of nursing control are much needed." Nonetheless, at the practical level, results showed extremely low values in terms of being "capable of": 1) judging the reliability of obtained information; 2) processing data using statistical software; 3) disseminating information while devoting attention to copyrights; and 4) understanding noteworthy points related to the use of ICT and knowing how to adapt in cases of trouble. In addition, those who had received information education were only 17.9%, which was found to be one factor preventing nursing managers from improving their practical capabilities of information control and use.

Keywords: clinical nursing, Nurse Administrator, information control practical use capability, e-learning

1 Introduction

Nurse Administrators at clinical nursing sites are key members of the fundamental organizational unit at a hospital. They strongly influence nursing and medical care quality depending on their control capability (e.g., people, things, money, information, and time). In recent years, middle nursing managers have been questioned about their control capabilities for education for new nurses, planning of continuing education, and for implementing new methods. "There is a gap separating educational sites and clinical education in terms of practical nursing capability of new nurses, the obligation to make efforts to provide training" has been placed on education for new nurses at clinical sites since 2010 "to acquire educational content steadily, particularly in clinical sites."[1] Figure 1 shows general human resource development in the nursing department of a hospital facility.

Fig 2 shows the present circumstances, under which middle nursing managers of a hospital's fundamental organizational unit process various information related to daily affairs. Middle managers of an organizational unit must process various information on a daily basis.

They are strongly required to have the capability of information control and use, which includes methods used to gather, organize, analyze, activate, and disseminate large amounts of information. Nursing managers

are also strongly required to have capabilities of information control and use, such as acquiring knowledge of information ethics. An earlier study evaluating the management capabilities of nursing managers revealed "information directivity is needed" among the items with low values of evaluated capability [2] . Additionally, in a study of "the competence of middle managers in a time of great change", reportedly all research collaborators (107 certified nurse administrators) considered that the capability of information control (gathering, grasping, sharing, analyzing, disseminating, providing, and using) was needed [3] . However, it is assumed that few middle managers who are responsible for information control at these sites have received information education. Therefore, ICT is inadequately applied. Middle managers must have or attain practical capabilities of information control and use, which includes acquiring knowledge of information ethics, as well as methods used to gather, organize, analyze, activate, and disseminate large amounts of information. Particularly in recent years, the introduction of ICT, including electronic medical records, continues to progress rapidly. Nonetheless, few Nurse Administrators who actually practice information control at these sites have received information education.

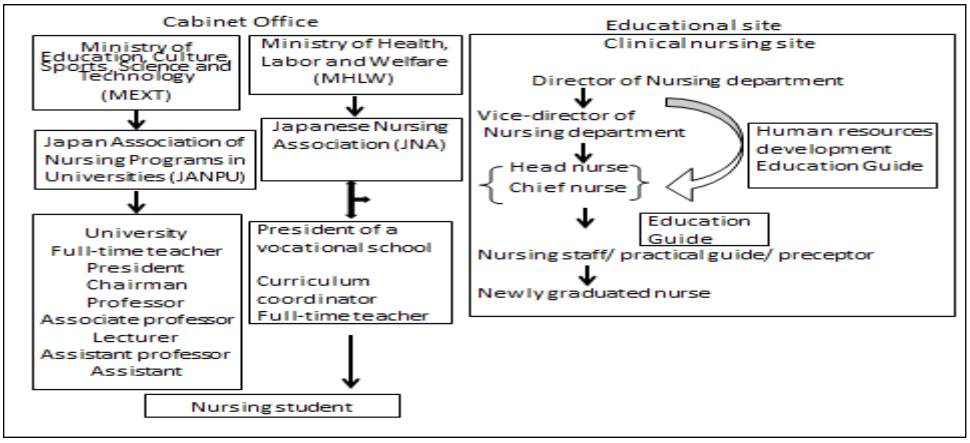


Fig1 Diagram of human resource development in the nursing department of a hospital facility.

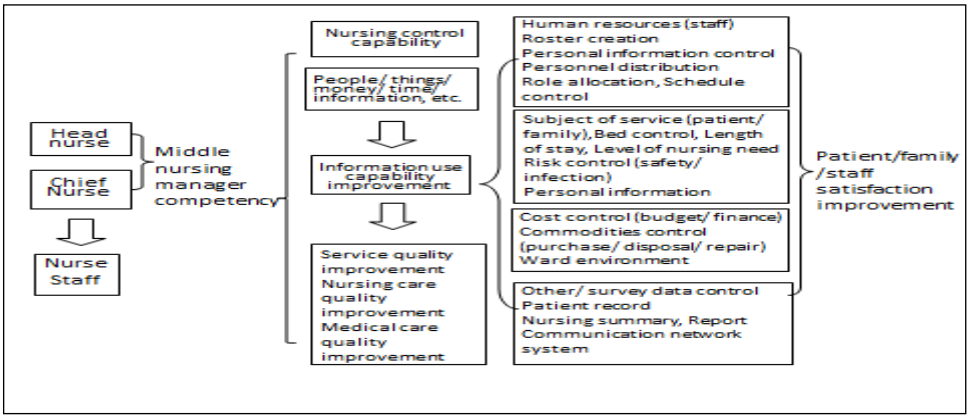


Fig2 Information control by Nurse Administrator in an organizational unit.

In the United States, the American Nursing Association (ANA) reported in 1994 that all nursing practitioners have ICT-related capabilities. Nursing informatics was recognized as a specialized area of nursing practice. However, recognition of the field of nursing informatics is notably low in Japan. Information education was included in official guidelines for school teaching in 1998. In 2004, information subjects were introduced into high schools. Considering those facts, many managers probably belong to the generation of students who did not receive information education in a systematic manner. An example of information education in the field of clinical nursing is the information control theory in certified nurse administrator training hosted by prefectural nurse centers. It set learning goals: "being capable of understanding the conceptualization of information and using it for nursing control" at the first level and "being capable of analyzing the situations of the introduction and using information technology in insurance medical welfare and looking to the future" at the second level. This report describes the results and problems of "a survey of the practical capability of information control and use of Nurse Administrators", which was conducted among people who took certified nurse Administrators training hosted by Japanese Nursing Association during August 2011 – December 2011.

2 Materials and Methods

We conducted a self-administered questionnaire survey of the practical capabilities of information control and use. Subjects were 116 people who took certified nurse administrator training held at Japanese Nursing Association during August 2011 – December 2011. Specific contents of the survey were 13 questions related to attributes and 16 questions related to the capabilities of information control and use. The survey employed a 5-point Likert scale, with responses ranging from 1 = "not at all needed" to 5 = "very much needed", and from 1 = "not at all capable" to 5 = "very capable". For explanation of the questionnaire, the researchers went to an actual training site and explained the aim and voluntary characteristics of the study. The analysis of results was based on the grand total. Centering on the correlation, we used statistical analysis software (SPSS19.0J; IBM Japan Ltd.), with a significance level of 5%. Regarding ethical considerations, we explained the confidentiality, the voluntary nature of research cooperation, the freedom to stop participation at any time, and the publication of results in writing. Consent was inferred when participants returned the questionnaire. This study was conducted under the approval of the ethics committee of Morinomiya University of Medical Sciences.

3 Results

The number of questionnaires collected was 116 (collection rate 100%). The nursing managers' positions were: 70 (60.3%) Head nurses, 32 (27.6%) Chief nurses, 2 (1.7%) Directors of nursing departments, 6 (5.2%) Vice-directors of Nursing departments, and 6 (5.2%) offering no response. The average age (mean±SD) was 44.1±5.3 years old. The average years of nursing experience were 22.0±5.7 years. Regarding the place of work, 83 (71.5%) worked in wards, 11 (9.5%) in outpatient units, 8 (6.9%) in operation rooms, 7 (6.0%) in nursing control offices, and 5 (4.3%) in other units. In terms of final specialized educational background, 1 (0.9%) had graduated from a graduate school, 10 (10.6%) from universities (including junior colleges), and 98 (84.4%) from nursing schools. To the question of "Have you ever received information education?" 23 (19.8%) answered "Yes" and 93 (80.2%) answered "No". To the question of "Do you like using computers?" 73 (62.9%) answered somewhere between "like and like somewhat" and

42 (36.2%) answered somewhere between "do not like very much and dislike". To the question of "Are information systems, such as electronic medical records, currently installed at your workplace?" 75 (64.7%) answered "installed", 26 (22.4%) "Planned for future installation", and 11 (9.5%) answered "not installed". Results (Mean±SD) from questions related to information control and use are presented respectively in Figure 3 and Figure 4.

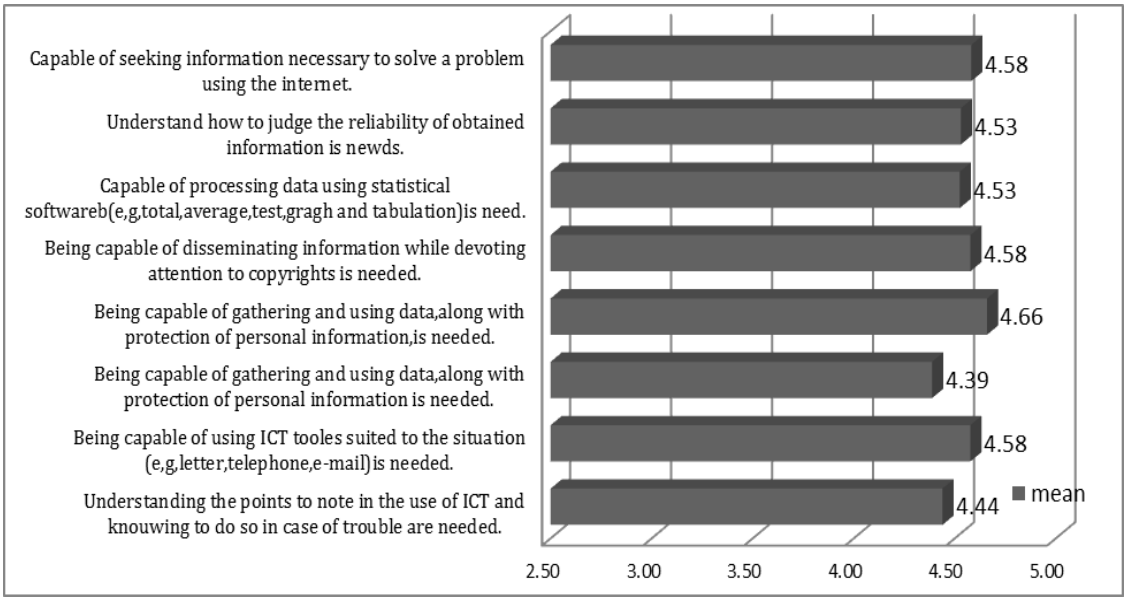


Fig 3 Capability of nursing information control and use.

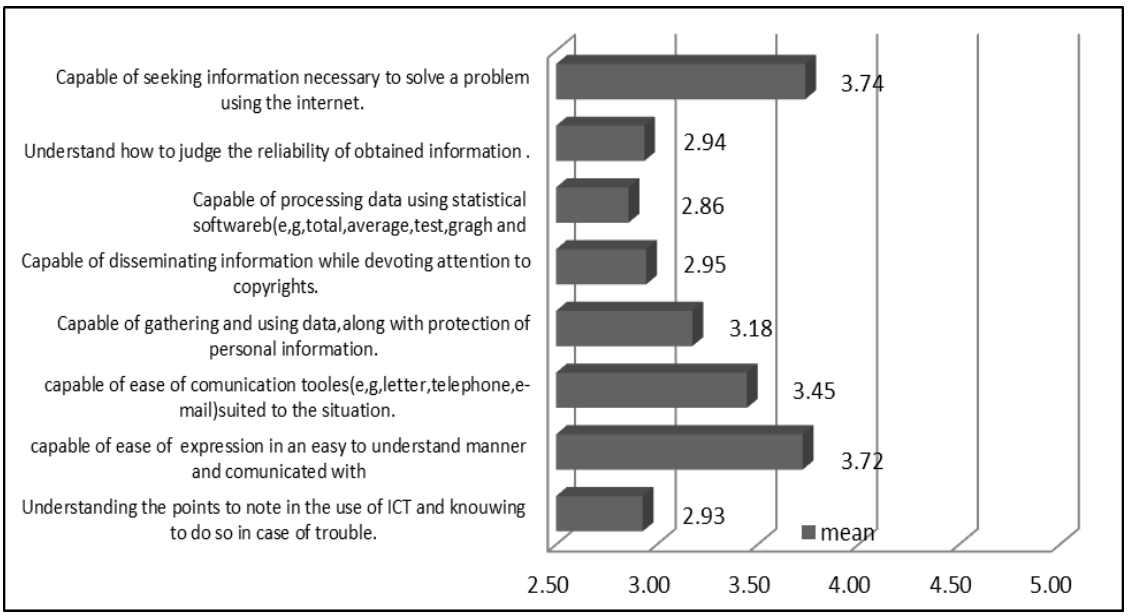


Fig 4 Practical capability of information control and use Capability.

In terms of the practical capability of nursing information control and use required for nursing managers, the values of results were 4.0 or greater for all questions. Particularly "being capable of gathering and using data, along with protecting personal information, is needed" showed the highest value of 4.66±0.71.

With regard to the practical capability of nursing information control and use, "capable of seeking information necessary to solve a problem using the internet" was 3.74±0.81, and "capable of expressing in an easy-to-understand way and communicating with other people using appropriate ICT tools" was 3.72±0.81. The result showing the lowest value at the practical level was "capable of processing data using statistical software (e.g., total, average, test, graph, and tabulation)" with 2.86±1.01 with subsequent "understanding how to judge the reliability of obtained information" with 2.94±0.899, "capable of disseminating information while devoting attention to copyrights" with 2.95±0.981, and "understanding the points to note in the use of ICT and knowing how to do so in case of trouble" with 2.93±0.896 (Figure4).

Next, as for the practical capability of nursing information control and use required for nursing managers, significant differences were found between "position". The fact might invariably demand information acquisition. In terms of the practical capability of nursing information control, significant differences exist between "Do you like using computers?" and "capable of disseminating information while devoting attention to copyrights", "understanding how to judge the reliability of obtained information".

4 Discussion and Conclusion

We conducted a survey of the practical capabilities of nursing information control and use among 116 people taking certified nurse administrator training. Results revealed that the average age was 44.1±5.3 years old and that the average years of nursing experience were 22±5.7 years. Regarding the experience of information education, 80.2% of the subjects "(had) not received information education." To the question of "Do you like using computers? 36.2% answered somewhere between "do not like very much and dislike". In addition, 31.9% of the facilities in which subjects worked did not introduce electronic medical records.

In Japan, information education was included in the official guidelines for school teaching in 1998. In 2004, information-technology related subjects were introduced into high schools. Considering those facts, many managers probably belong to the generation of students who did not receive information education in a systematic manner. Currently, information control theory is taught in certified nurse administrator training. However, an annual quota exists for people who can take training. Nursing managers have few opportunities to receive information education. These circumstances might influence the practical capability of information use at clinical sites.

Next, in terms of the practical capability of nursing information control and use required for nursing managers, the values of results were 4.0 or greater for all questions. Particularly "being capable of gathering and using data, along with protecting personal information, is needed" showed a high value of 4.66±0.71. In recent years, because of the establishment of more nursing universities and graduate schools, guides and guiding systems for clinical research are regarded as being ready. Clinical research has become active. Opportunities to receive information education are also increasing: environments capable of yielding research literature and information directly are increasing, resulting from the installation and the improvement of document retrieval systems. In terms of the fact that "capable of processing data using statistical software (e.g., total, average, test, graph, and tabulation)" showed a value of

2.86±1.01, learning of specific statistical methods might not be sufficient. Moreover, being unable to use statistical methods might not create any particular disadvantage. Nevertheless, nursing managers who control large amounts of information in wards, if they could use that information, might thereby improve the quality of nursing care. The service level and the satisfaction of patients and families would improve, and staff would increase in the future.

Moreover, regarding the practical capability of information control and use required for nursing managers, a marked difference was found between “position” and “seeking information necessary to solve a problem using the internet is needed.” The higher the employee position, the more tasks they would have to handle when trouble arises. This fact might invariably demand information acquisition. Significant differences were found between “the number of beds and capability of processing data using statistical software”. This result was attributed to the following reasons: of the subjects in this survey, 83 (71.5%) were working in wards and those involved in controlling large amounts of information, as shown in Figure 2. In terms of the practical capability of nursing information control, significant differences were found between “Do you like using computers?” and “capable of disseminating information while devoting attention to copyrights” and between “with or without experience of information education and age and years of service”. The more opportunities employees have to use information equipment, the more knowledge of operation and systems increase. Consequently, employees are expected to be able to overcome diffidence. As indicated by the discussion presented above, using the practical capability of information control and using it at actual sites in the future, Strategies for overcoming diffidence to information equipment are necessary. Nursing managers have time constraints both in their private and professional lives. To encourage them to use computers and the internet, their environment must be improved because they can contribute immediately. One example is the installation of computers connected to the internet in each ward. In addition to that access, a system to inform employees how to operate and use the computers must be developed.

References

[1]The e-Nurse Center, the Japanese Nursing Association: A Report on Early Unem Ployment among Newly Graduated Nurses in Fiscal Year 2004; 2005.

[2]Shigeko Yokoyama, Rie Yokoyama, et.al.: Research [on a nursing administration person's management ability factor], and Sophia University science-of-nursing part bulletin, Vol.2(2), pp.59-65 (2005)

[3]Kotomi Matsuda: Competency of the middle manager of a change term, the Shiga University of Medical Science graduate school master's thesis, pp. (2008)

[4]Takako Izumi, Yukie Majima, Satoshi Shimada :The reality and challenges of middle nursing manager participation in the e-learning for new graduated nurses, International Proceedings of Economics Development and Rsearch, Education and Management Innovaiton II, Vol.60(21), pp.102-106 (2013)

[5]Akemi Hirata: Recognition about the administrator role in the Certified Nurse Administrator education course first level traineeds, Yokohama journal of Nursing Vol.6(1), pp.35-40 (2013)

[6]Manami Inoue, et.al.: The actual situation and the problem of the Certified Nurse Administrator’s education in yamaguchi prefecture, Journal of yamaguchi prefecture universty,Vol.pp.43-47 (2013)

[7] Kimiko Kathuyama: What will be the next generation program for nurse managers in Japan? -Leaning from the U.S.programs-, Journal School of Nursing, Osaka Prefecture University, Vol.16(1), pp.(2010)

A Benchmarking Study of Standard Frameworks for Information Technology Governance

Wilai Sunthonwutinun† **Vasin Chooprayoon‡***

†‡*Faculty of Information Technology, Rangsit University, Paholyothin Rd., Muang-ake, Lak-hok, Muang, Pathumthani 12000 Thailand
wilai55@rsu.ac.th; vasin@rsu.ac.th

Abstract

Benchmarking the standard frameworks in market place is one of important approaches for selecting appropriate standard frameworks used in information technology (IT) governance in order to balance requirements of stakeholders, risks, and related resources to achieve the organization goal. This study proposes different characteristics of each framework used for IT governance. The findings showed that each framework has different specific characteristics which involve with IT governance, there are six requirements which are covered by the eight standard frameworks: ISO/IEC 38500:2008, ITIL®V3 2011, ISO/IEC 27000 Series, ISO/IEC 31000 Series, TOGAF®, CMMI, PRINCE2.and CobiT5. The six requirements are 1) management of business risk, 2) optimization of business process function, 3) operational and staff productivity, 4) skilled and motivated people, 5) process improvement, and 6) IT management matching with business requirements. Cobit5 and ITIL V3 2011 frameworks compose of the most requirements of business IT governance (26 and 22 requirements in ascending order).

Keywords: IT Governance; Benchmarking of IT Governance Standard Frameworks

1 Introduction

IT utilization and a business goal of an organization is a foundation of efficient and effective IT governance [1]. Over the past decade, “good governance” has become a key issue in business thought. A company board and executives need to pay intention to IT as an important tool for business performance. Integration between information and IT has scattered over enterprises, communities, and public sectors. The integration plays significant function in business environment. Most enterprises attempt to organize high quality of information and IT readiness in order to support decision makings in business areas. IT investment is a track for generating business value. Setting up IT investment as a strategic goal of enterprises lead to gain more business benefits and achieve operational excellence. It is reasonable to point out that maintaining IT related risk at an acceptable level, optimizing cost of IT services and technology, complying with laws, regulations, contractual agreements and policy are good performances for handling IT governance [2].

Most enterprises need standard tools to assist their executives to perform good IT governance including the IT best practice management. The previous studies showed that there were various standard tools used in corporate activities concerning with information technology [3]. (IEEE "Eye-triple-E" 2013) In market places, there are varieties of standard frameworks: Cobit5, ITIL, ISO/IEC involved IT, PRINCE2, TOGAF, CMMI, etc. The enterprises intricately choose the appropriate frameworks because they have to spend long time for studying all standards in market places. A benchmark of each standard is an important key approach for making decision on selecting the standard tools for IT governance.

1.1 IT Governance

IT Governance is an approach that enterprises arrange IT strategies within their business to ensure achievement with regards to the strategies including implementation of appropriate methods to evaluate IT performance [4].

The IT governance concepts in each enterprise are based on balancing of benefits, risk levels, resource utilization, and holistic management [5]. In addition, the IT governance involves with management principles of information and communication technology (ICT) in organizations emphasizing IT values with regarding to organization strategies. The IT governance is a part of corporate governance which is composed of accountability and fiduciary duty, economic efficiency, strategic efficiency, and stakeholders' roles. In summary, IT governance concentrates on effectiveness of IT performance and IT risks in organizations [6]

1.2 Various standard frameworks for IT governance

ISO/IEC 38500:2008 is a corporate governance standard applied to the IT governance. The standard is composed of six principles: 1) responsibility, 2) strategy, 3) acquisition, 4) performance, 5) conformance, and 6) human behaviors. The executives use the principles for governing their organizations [7].

ITIL®V3 2011 is an IT service management standard based on five principles: 1) service strategy, 2) service design, 3) service transition, 4) service operation, and 5) continual service improvement. The principles explain the best practices in IT service management. The principles concentrate on measurement and improvement of the quality of IT services continually [8].

ISO/IEC 20000 is a part of **ITIL®V3** focusing on IT service management checklist in order to ensure that the service management has been performed correctly. [9].

ISO/IEC 27000 Series emphasizes how IT security techniques and information security management system govern IT performance in organizations. The series based on the codes of practices on information security management, governance objectives, and range of specific security controls [10].

ISO/IEC 31000 Series provides principles and guidelines for risk management in organizations. The principles aim to be applied to any activities, processes, functions, projects, products, services and assets including monitoring, reviewing, recording and reporting the results [11].

TOGAF® provides enterprise architecture (EA) methodology utilizes resources and realizes return on investment. The context of EA is able to define as a system or a system plan. TOGAF supports four architecture types: 1) business architecture, 2) data architecture, 3) application architecture, and 4) technology architecture [12].

CMMI is a standard process used in performance management. CMMI is composed of five levels: 1) initial level; 2) managed level which consisted of seven process areas--requirement management, project planning, project monitoring and control, supplier agreement management, measurement and analysis, process and product quality assurance, and configuration management; 3) defined level which consisted of eleven process areas--requirements development, technical solution, product integration, verification, validation, risk management, integrated project management, organizational process definition, organizational process focus, organizational training, and decision analysis and resolution; 4) quantitatively managed level which composed of two process areas--organizational process performance and quantitative project management; and 5) optimizing level which consisted of two process areas--causal analysis and resolution and organizational performance management [13T.

PRINCE2 is a process-based method used for governing project management. The key features of the method are a) business justification,
b) defining organization structure for project management team, c) product-based planning,
d) dividing the project into manageable and controllable stages, and e) flexibility [14].

CobiT5 is a business framework for IT governance and management of enterprises. Cobit5 provides five principles and seven enablers to maximize the value and minimize the risk of enterprises. The five principles are 1) meeting stakeholder needs, 2) covering the enterprise end to end, 3) applying a single integrated framework, 4) Enabling a holistic approach, and 5) Separating governance from management. The seven enablers are
1) principles, policies and frameworks, 2)

processes, 3) organizational structures, 4) culture, ethics and behavior, 5) information, 6) services, infrastructure and application, and 7) people, skill and competencies [2].

1.3 Level of participation on business IT governance with the frameworks

The study from PricewaterhouseCoopers [16] pointed out that fifty CIOs from worldwide enterprises have accepted the frameworks and used them as a mechanic for running best practices in their organisations. Almost CIOs (95%) have followed the well-known frameworks' guidelines such as CobiT (63%) and ITIL (60%). The study also showed that the frameworks are separately used according to requirements of the enterprises. There was least combination between CobiT and ITIL.

2 Objective of the Study

The objective of this study is to compare each standard framework in terms of different and similar characteristics matching with IT governance. This study concentrates on eight standard frameworks involve with IT governance as shown in Figure 1.

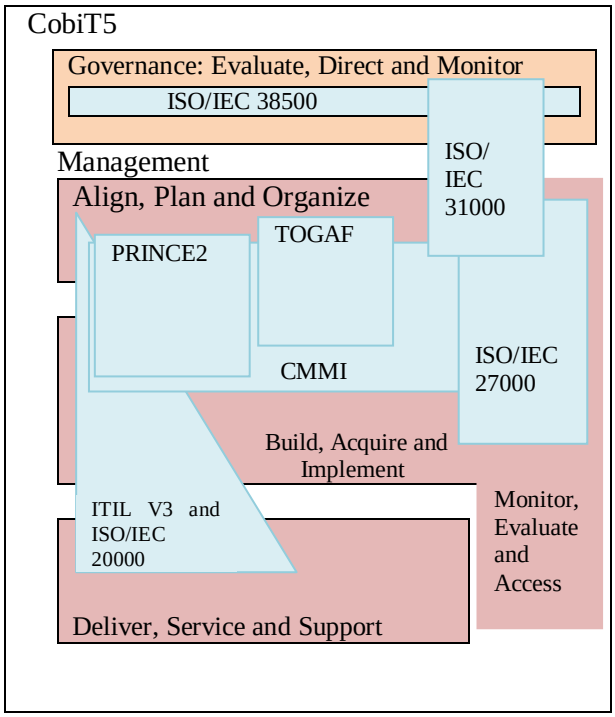


Figure 1 Standard Frameworks (Adapted from [2])

3 Scope of the Study

The scope of this study is based on eight standard frameworks: ISO/IEC 38500:2008, ITIL®V3 2011, ISO/IEC 27000 Series, ISO/IEC 31000 Series, TOGAF®, CMMI, PRINCE2.and CobiT5

4 Study Framework

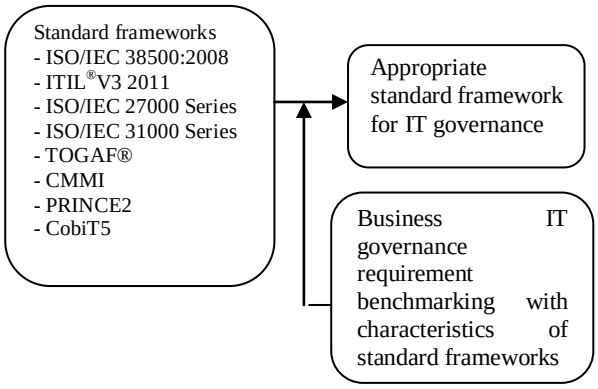


Figure 2 Study framework

This study conducted with regarding to the framework showed in Figure 2 which composed of eight standards. The standards lead to appropriate framework for IT governance which derives from requirements of the enterprises on characteristics of IT governance. Table 1 shows a comparison study of IT governance characteristics appeared in each standard.

Table 1. Characteristic of IT Governance frameworks

Characteristic of IT Governance framework	Standard frameworks							
			ISO/IEC					
	CobiT5	ITIL®V3 2011	38500:2008	27000 Series	31000 Series	TOGAF®	CMMI	PRINCE2
1. IT Governance and management (GM)	X							
2. Corporate governance of information technology (CG)			X					
3. IT service management (SM)		X						
4. IT security technique,				X				

Characteristic of IT Governance framework	Standard frameworks							
	CobiT5	ITIL®V3 2011	ISO/IEC 38500:2008	ISO/IEC 27000 Series	ISO/IEC 31000 Series	TOGAF®	CMMI	PRINCE2
Information security management system (ISMS)								
5. IT risk management (RM)					X			
6. Enterprise architecture methodology (EA)						X		
7. Performance management (PM)							X	
8. Project management (PJM)								X

In terms of business IT governance requirements, there are three parts: 1) direct process based on BSC (Balance Score Card) and stakeholders goal based on CobiT5 parameter; 2) evaluate process based on performance management aligning direction of IT with IT enterprise goal, value delivery, risk management, resource management; and 3) monitoring process based on satisfaction, improve processes, internal control, independent assurance and audit. Table 2 shows the requirements in details.

Table 2. Business IT Governance Requirements

Governance Process	Adapted from
1) Direct (Enterprise Goal base on BSC)	BSC [2]
a) Financial	
1. Stakeholder value of business investment	
2. Portfolio of competitive products and services	
3. Managed business risk	
4. Compliance with external laws and regulations	
5. Financial transparency	
b) Customer	
1. Customer - oriented service culture	
2. Business service continuity and availability	
3. Agile responses to a changing business environments	
4. Information- based strategic decision making	
5. Optimization of service delivery costs	
c) Internal	
1. Optimization of business process function	
2. Optimization of business process costs	
3. Managed business change programmed	
4. Operational and staff productivity	
5. Compliance with internal policy	
d) Learning and Growth	
1. Skilled and motivated people	
2. Product and business innovation culture	
2) Evaluate (Performance measurement)	
1. Alignment of strategic direction of IT with enterprise goal	Researcher defined
2. Value delivery of IT or return on investment (ROI)	National Computing Center [15]
3. Risk management	
4.Resource management use of IT resources, IT capability and infrastructure compared to current and expected future business requirements	
3) Monitor	
1. Satisfaction	Researcher defined
2. Improve processes	National Computing Center [15]
3. Assess Internal Control Adequacy	
4. Obtain Independent Assurance	
5. Provide for Independent Audit	

5. Methodology

This study is a qualitative research using content analysis method for benchmarking standard frameworks on IT governance. The steps of the analysis are as follows:

1) Review literature on potential IT governance frameworks,
2) Classify each different characteristics of IT governance frameworks,
3) Analyze business IT governance requirements,
4) Compare Characteristics of each framework with business IT governance requirements, and
5) Evaluate and choose the appropriate

framework for IT governance.

6. Research Findings

The benchmarking resulted that there are six requirements which are covered by the eight standard frameworks: 1) management of business risk, 2) optimization of business process function, 3) operational and staff productivity, 4) skilled and motivated people, 5) process improvement, and 6) IT management matching with business requirements (Table 3). Cobit5 and ITIL V3 2011 frameworks compose of the most requirements of business IT governance (26 and 22 requirements in ascending order). The findings bring to the development of the proposed appropriate standard framework for IT governance enterprises as shown in Table 4 which points out how each standard framework is suitable for each requirement. In addition [3], eleven selected papers related with the frameworks were studied; the study showed that five papers focused on CobiT or ITIL and six papers focused on various framework combination such CobiT, ITIL, ISO/IEC, CMMI and Prince2.

7. Conclusions

Implication of the findings to the enterprises is based on their requirements and effectiveness of their evaluation and assessment. The proposed model of appropriate framework for business IT governance (Table 4) is a roadmap for choosing a standard framework. This study anticipates the proposed model to play an important role in terms of contribution to decision making on how to select standard frameworks for IT governance.

Table 3 Benchmarking model between business IT governance requirements with characteristics of the standard frameworks

Business IT Governance Requirements	Characteristic of Standard Frameworks							
Governance Process	CobiT5	ITIL® V3 2011	ISO/IEC			TOGAF®	CMMI	PRINCE 2
			38500: 2008	27000 Series	31000 Series			
	GM	SM	CG	ISMS	RM	EA	PM	PJM
1) Direct (Enterprise Goal base on BSC)								
a) Financial								
1. Stakeholder value of business investment	X	X	X	X	X	X	X	
2. Portfolio of competitive products and services	X	X			X	X	X	X
3. Managed business risk	X	X	X	X	X	X	X	X
4. Compliance with external laws and regulations	X	X	X	X	X	X		
5. Financial transparency	X		X		X			
b) Customer								
1. Customer - oriented service culture	X	X			X			
2. Business service continuity and availability	X	X	X	X	X	X	X	
3. Agile responses to a changing business environment	X	X			X			X
4. Information- based strategic decision making	X	X	X	X	X	X		X
5. Optimization of service delivery costs	X	X	X		X	X	X	X
c) Internal								
1. Optimization of business process function	X	X	X	X	X	X	X	X
2. Optimization of business process costs	X	X	X		X	X	X	X
3. Managed business change programs	X	X			X			X
4. Operational and staff productivity	X	X	X	X	X	X	X	X
5. Compliance with internal policy	X	X	X	X	X	X		
d) Learning and Growth								

Business IT Governance Requirements	Characteristic of Standard Frameworks							
Governance Process	CobiT5	ITIL® V3 2011	ISO/IEC 38500: 2008	ISO/IEC 27000 Series	ISO/IEC 31000 Series	TOGAF®	CMMI	PRINCE 2
	GM	SM	CG	ISMS	RM	EA	PM	PJM
1. Skilled and motivated people	X	X	X	X	X	X	X	X
2. Product and business innovation culture	X		X			X	X	
2) Evaluate (Performance measurement)								
1. Alignment of strategic direction of IT with enterprise goal	X	X	X	X	X		X	X
2. Value delivery of IT or investment (ROI)	X	X		X			X	X
3. Risk management	X	X		X	X		X	X
4. Resource management use of IT resources, IT capability and infrastructure compared to current and expected future business requirements	X	X	X	X	X	X	X	X
3) Monitor								
1. Satisfaction	X	X		X				X
2. Improve processes	X	X	X	X	X	X	X	X
3. Assess Internal Control Adequacy	X	X	X	X	X			
4. Obtain Independent Assurance	X						X	
5. Provide for Independent Audit	X			X				
Total	26	22	17	17	21	15	16	16

Table 4 Appropriate standard framework for IT governance in the enterprises

Governance Process Requirements		Appropriate Standard Frameworks							
Direct	1. Stakeholder value of business investment	GM	SM	CG	ISMS	RM	EA	PM	PJM
	2. Portfolio of competitive products and services								
	3. Managed business risk			CG	ISMS				
	4. Compliance with external laws and regulations								
	5. Business service continuity and availability							PM	
	6. Information- based strategic decision making								PJM
	7. Optimization of service delivery costs							PM	
	8. Optimization of business process function				ISMS				
	9. Optimization of business process costs								
	11. Operational and staff productivity			CG	ISMS				
	12. Compliance with internal policy								
	13. Skilled and motivated people							PM	PJM
Evaluate	1. Alignment of strategic direction of IT with enterprise goal								
	2. Value delivery of IT or investment (ROI)								
	3. Risk management					RM			
	4. Resource management use of IT resources, IT capability and infrastructure compared to current and expected future business requirements			CG			EA		
Monitor	1. Improve processes								
	2. Assess Internal Control Adequacy								

References

[1] Syed Jamal Hussain Muhammad Sibghatullah Siddiqui. 2005. "Quantified Model of CobiT for Corpoetate IT Governance." (Accessed 16 May, 2013).

[2] ISACA. 2013. "Cobit5" In A business framework for the governance and management of enterprise IT. 13. Original edition, Printed in the United States of America.

[3] IEEE "Eye-triple-E". 2013. "Search Article Full Text." http://ieeexplore.ieee.org/search/searchresult.jsp?newsearch=true&queryText=IT Governance Standard (24 May 2013).

[4]"IT Governance Definition and Solutions". 2013.CIO.com. http://www.cio.com/article/111700/IT_Governance_Definition_and_Solutions#what

[5] Michael Holm Larsen; Mogens Kuhn Pedersen; Kim Viborg Andersen. 2006. "IT Governance: Reviewing 17 IT Governance Tools and Analysing the case of Novozymes A/S." (2006).

[6] Z. Shadid, G. Boyd and R. Brisebois, "What Is IT Governance? And Why Is It Important for the IS Auditor?" INTO IT, Toronto, 2008.

[7] ISO (International Organization Standardization), for. 2008. "ISO/IEC 38500:2008 Corporate governance of information technology." http://www.iso.org/iso/catalogue_detail?csnumber=51639 (19 May 2013).

[8] Alison cartidge :Ash Hanna: Colin Rudd : Ivor macfarlane : John Windebank : Suart Rance. 2007. An Introduction Overview of ITIL@V3. Version1.0 ed.: The UK Chapter of the itSMS.

[9] ISO (International Organization for Standardization). 2012. "ISO/IEC 20000-3:2012 Information technology Service management ". http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=60031 (19 May 2013).

[10] ISO (International Organization for Standardization).2012."ISO/IEC27000 Information technology Security techniques Information security management systems." http://standards.iso.org/ittf/licence.html (19 May 2013).

[11] ISO (International Organization for Standardization). 2008. "ISO/IEC31000 Risk management Principles and guidelines on implementation." http://standards.iso.org/ittf/licence.html (21 May 2013).

[12] TOGAF (The Open Group Architecture Forum):Andrew Josey. "TOGAF™ Version 9 Enterprise Edition An Introduction." In. The Open Group 44 Montgomery St. #960 San Francisco, CA 94104: The Open Group, January 2009.

[13] Capability Maturity Model® Integration (CMMI). 2005. Capability Maturity Model® Integration (CMMI®) Overview." http://cmmiinstitute.com/ (21 May 2013).

[14] ILX Group; the Cabinet Office. 2013. "PRINCE2®, What is PRINCE2? PRINCE2 Definition." http://www.prince2.com/what-is-prince2.asp (21 May 2013).

[15] National Computing Center. "IT Governance Developing a successful governance strategy A Best Practice guide for decision makers in IT." www.ncc.co.uk (24 June 2013).

[16] PricewaterhouseCoopers. 2007. "IT Governance in Practice – Insights from leading CIOs." www.pwc.com (6 August 2013).

Thai ICT HR Competency Standard Conceptual Framework for Supporting Sustainable Smart Thailand Towards AEC 2015

Kriengkrai Bhuvanij‡ M.L.Kulthon Kasemsan* Prasong Praneetpolgrang**

‡Faculty of Information Technology, Rangsit University Thailand
kriengkr@th.ibm.com

*Faculty of Information Technology, Rangsit University, Thailand
kkasemsan@rsu.ac.th

**Information Science Institute, Sripatum University, Thailand
prasong.pr@spu.ac.th

Abstract

This paper presents a Competency Standard Framework for Thai ICT human resource to support a sustainable Smart Thailand policy which will promote Thailand growth to be well balanced in the development of economic, environment and social areas. The Framework will eventually pave the way for Thailand's journey towards knowledge economy and ASEAN integration.

Keywords: ICT Competency, Smart Thailand, Sustainable Development, AEC2015.

1. Introduction

As a major driving force of regional integration, the ASEAN Economic Community (AEC) 2015 is part of a larger plan for the ASEAN Community, which involves security, economic, and socio-cultural cooperation. AEC 2015 aims to transform ASEAN into a single market and production base, a highly competitive economic region and a region fully integrated into the global economy. To this end, the AEC blueprint focuses on free flow of goods, services, investment, and skilled labor, and freer flows of capital, to help minimize transaction costs and maximize trade gains [1].

At the 13th ASEAN Summit on 20 November 2007, the ASEAN countries aimed to enhance concerted and collective cooperation in building of the Information Society and to increase the region's connectivity and competitiveness, strengthening on human resource cooperation, particularly on regional standardization of ICT human capital competencies and regional recognition across ASEAN Member States of ICT skills certification to increase knowledgeable and versatile workforce in efforts to build an inclusive knowledge-based ASEAN Community in 2015 [2].

More importantly, after the success of its economy integration in 2015, ASEAN countries will need to address its sustainable development challenges and priorities for the region over the next decade. The ASEAN Vision 2020 provides the long-term sustainable development framework for the sub-region and its sustainable development priorities for ASEAN would be useful for the government, non-government, private, regional and international organizations in the pursuit of developing policies, strategies, and action plans. One of the important prerequisites to the pursuit of the subregion's sustainable development goals is the Co-ordination and integration at the Southeast Asian subregional level of economic, social, and environmental initiatives and policies [3].

With the long-term aim towards ASEAN Community level, Thailand has been paving its way to prepare the country for ASEAN Community which knowledgeable and versatile workforce will be required through the Thailand ICT Policy Framework (2001-2010) or IT2010 and ICT2020 Policy Framework aiming to enhance the economy and quality of life of the

Thai people and lead Thailand towards a knowledge-based economy and society by using ICT as a driving force for creating knowledge, creativity and innovation in goods and services.

Under the ICT 2020 Policy Framework, Thailand envisions for smart development, with a knowledge- and wisdom-based economy and society. The "Smart Thailand 2020" vision as depicted in Fig.1 states that ICT is a key driving force in leading Thai people towards knowledge and wisdom and leading society towards equality, stronger economy and environment sustainability [4].

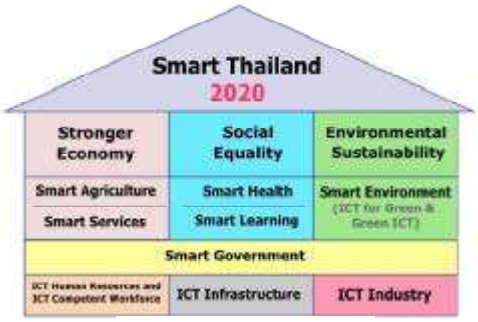


Fig.1 Thailand ICT 2020 Framework [4]

Severino [5] noted that an assessment is being undertaken of the needs of ASEAN countries, including Thailand, in terms of human resources for ICT. They are cooperating in the training of their human resources, which is the key to ICT and the knowledge economy, the economy of the twenty-first century. Information and communications technology is thus another critical element in the integration of the ASEAN economy and of the region as a whole.

Since both ICT and ICT human resource are the two major elements to drive Thailand its goals in 2020, particularly on the ICT human resource, hence the national ICT skills framework for human resource will be required and needs to be assessed to possibly help the country achieve sustainable Smart Thailand and ASEAN Economic Community in 2015.

In this paper, the Thai HR ICT Competency Standard Framework for supporting sustainable Smart Thailand will be explored.

2. Theoretical Background

2.1 Defining Sustainability and Sustainable Development

The word "sustain" derives from the Latin "sustenere" [6]. In its primary sense it can be described as survival assurance meaning that an economical, ecological or social system should be preserved for future generations and, thus, necessary resources should only be exploited to a degree where it is possible to restore them within a regeneration cycle. The most common definition from the Brundtland Commission defines sustainability as a "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [6], [7]. All definitions of sustainability have the preservation of the economical, ecological and social system for the benefit of future generations in common. These dimensions represent the three main pillars of sustainability and are known as the "triple-bottom-line" concept [6], [8].

The "triple-bottom-line" concept provides a framework to companies to measure and report their performance and organizational success in relation to these pillars. Especially at the business level, sustainability is mainly equated with the economical or financial sustainability [6], [9]. Triple bottom line (TBL) accounting expands the traditional reporting framework to take into account social and environmental performance in addition to financial performance. In 1981, Freer Spreckley first articulated the triple bottom line in a publication called 'Social Audit - A Management Tool for Co-operative Working'. In this work, he argued that enterprises should measure and report on social, environmental and financial performance [10], [11].

Although the definition of sustainable development emerged from an international enquiry into the relationship between environment and development, it is not concerned primarily with the environment but with the sustainability of the overall developmental context. This usually comprises three main elements [12] as the followings:

1. Economic development – reducing and seeking to eradicate income poverty, achieving

higher levels of prosperity and enabling continued gains in economic welfare;

2. Social development – reducing and seeking to eradicate other dimensions of poverty, improving the quality of education, health, housing and other aspects of the welfare of individuals and communities, and enhancing the quality of social interaction, engagement and empowerment;

3. Environmental protection – reducing pollution and other negative impacts on the environment, mitigating the effects of industrialization and human activity, and seeking to achieve sustainable use of resources in the interest of future generations.

This relationship is sometimes illustrated either as a Venn diagram or through pillars as per Fig.2.1 and 2.2 below:

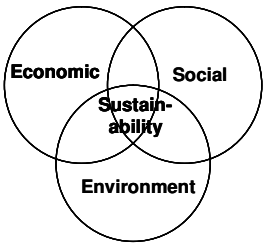


Fig.2.1 Venn diagram of Sustainable Development [12]

Fig.2.2 Three Pillars of Sustainable Development [13]

Three pillars or circles of "economic development, social development and environmental protection" as three "interdependent and mutually reinforcing pillars" of sustainable development [12], [13],[14].

2.2 Defining Competency and ICT Competency

2.2.1 Competency Ice Berg Model - Meaning and its Components [15]

The iceberg model for competencies takes the help of an iceberg to explain the concept of competency. An iceberg which has just one-ninth of its volume above water and the rest remains beneath the surface in the sea. Similarly, a competency has some components which are visible like knowledge and skills but other behavioural components like attitude, traits, thinking styles, self-image, organizational fit etc are hidden or beneath the surface.

2.2.2 Ice-Berg Model [15]

The pictorial representation of the Ice-Berg model of competency is as Fig.3 below:

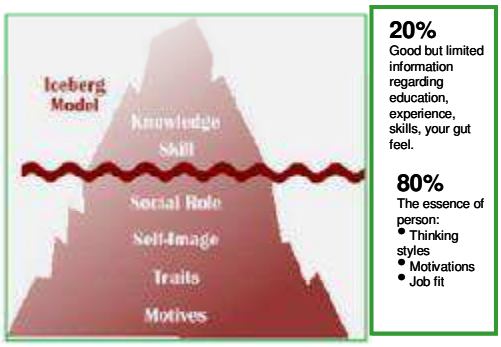


Fig.3 Competency Ice Berg Model [15]

Where the different components of the model can be described as Fig.4:

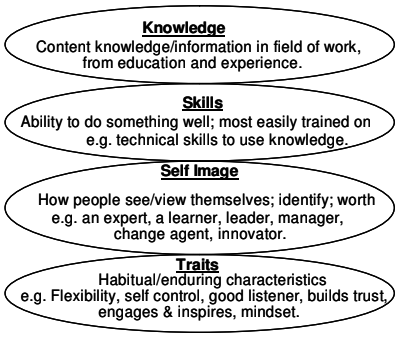


Fig.4 Components of Competency [15]

The aspects of competencies which lie below the surface like attitude, traits, thinking styles etc. directly influence the usage of knowledge and skills to complete a job effectively.

Developing the two levels of competencies also takes different routes. The visible

competencies like knowledge and skills can be easily developed through training and skill building exercises however the behavioral competencies are rather difficult to assess and develop. It takes more time and effort intensive exercises, like psychotherapy, counseling, coaching and mentoring, developmental experiences etc.

Besides that, OECD also defined competency as followings:

"A competency is more than just knowledge and skills. It involves the ability to meet complex demands, by drawing on and mobilizing psychosocial resources (including skills and attitudes) in a particular context. For example, the ability to communicate effectively is a competency that may draw on an individual's knowledge of language, practical IT skills and attitudes towards those with whom he or she is communicating"[16].

2.2.3 ICT Competency

Information technology (IT) competency may refer to many different knowledge areas and skill sets. Most often, the IT competency of a person refers to their own core competence or knowledge in IT systems. For a professional, the IT competency represents his own skill set with information systems and other technology. The professional generally needs to have a firm knowledge of current hardware, software and other technologies such as computer systems as well as the ability to quickly learn and adapt to the constant changes in technology, project management, planning and working within business practices. As such, ICT competency entails much more for a company as a whole than simply understanding current technology. While each individual may not need to have technical skills and knowledge, having enough people staffed to handle technical applications, people management and the business politics is crucial [17].

Hence, with the definition of competency and ICT competency, it could be concluded that the ICT competency could be categorized in the two layers i.e. Layer 1; the skills and knowledge on ICT including other related skills for ICT professional and Layer 2; the behavioural competency components such as attitude, traits, self image and thinking styles towards ICT and the use of ICT.

2.3 ICT Skills Framework

Nuangjamnong et al. [18] also outlined a common definition of skills framework architecture on knowledge, qualifications, skills performance standards and experience requirements against a complete set of roles and standard job specifications as in Fig.5.

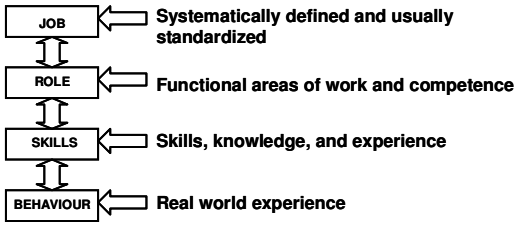


Fig.5 Components of Competency [18]

In addition, Nuangjamnong et al. [18] also explained the study of one of the well-recognized standard competency frameworks, which can be applied into real IT environment, on The Skills Framework for the Information Age (SFIA) which is a well-known IT standard skills framework and widely used in United Kingdom (UK). SFIA provides a general reference model, which identifies the skills needed to develop effective Information Systems (IS) making use of Information and Communication Technology (ICT) [19]. The framework integrates technical and academic skills, which are required by all employers and across all industries.

The main characteristics of SFIA in Human Resource (HR) activities are: skill audit (assessment / assignment / recruitment / external or internal). It can also be used for planning future skill requirement (skills gap analysis, and management of the corporate competency profiles); development of training courses (developing and maintaining a business-oriented ontology); standardisation of job titles and functions as well as resources allocation. The SFIA standard is intended to be used as an IT skills management tool, which may help data centres in Thailand to recruit suitable IT staff.

2.4 Skills Framework for the Information Age (SFIA)

The Skills Framework for the Information Age (SFIA) is a model for describing and managing competencies for ICT professionals for the 21st century, and is intended to help match the skills of the workforce to the needs of the business. It maps out the range of skills as a two-dimensional table, by tagging each skill with a category and responsibility level.

These categories are divided into six main areas: Strategy and planning; Business change; Solutions development and implementation; Service management; Procurement and management support; and Client interface. While on the second dimension, level of responsibility, there are seven levels, in ascending order: Follow; Assist; Apply; Enable; Ensure and advise; Initiate and influence; and Set strategy, inspire and mobilize. Each of these responsibility levels has a generic description showing the level of autonomy, influence, complexity, and business skills required [20].

SFIA's aim is to provide a management tool to help those who are the definitions provide precise statements of the various levels of skill required. The IT industry contains a wealth of information, formal and informal, supporting each skill. This covers many complex aspects, processes and methods that may relate to the skill. SFIA's purpose is not to include that information, but to provide a management tool that helps managers make sense of the complexity. SFIA's descriptors are not, in general, described in terms of technologies or products [21].

2.5 Skills for the 21st Century

Success in the 21st century requires the ability to access, synthesize, and communicate information; to work collaboratively across differences to solve complex problems; and to create new knowledge through the innovative use of multiple technologies. To meet demands of the new global economy, particularly towards ASEAN Integration and ASEAN Community, one could leverage the 21st century Life and Career Skills assist in synthesizing information, working effectively in diverse teams, managing complex projects, and demonstrating responsibility to the community and environment. The review of the literature on 21st Century Skills suggests that education must be upgraded for learners to thrive in the new global economy. In short, 21st Century Skills are more than technological literacy, instead they include proficiency in critical thinking, problem solving, communication, and teamwork [22].

2.6 The Thai government policy on Sustainable Smart Thailand

As per the speech of Minister of ICT of Thailand in December 2012 in the "CIO 16 Annual Conference: ICT Sustainability for Smart Thailand", the vision of Smart Thailand 2020 coincides and supports the ASEAN ICT Master Plan 2015 and ASEAN 2020 vision [23].

With the "Smart Thailand 2020" vision which states that ICT is a key driving force in leading Thai people towards knowledge and wisdom and leading society towards equality and sustainable economy, "the ICT Sustainability for Smart Thailand", as noted by the minister, will be aimed to drive for sustainable development in the balanced of the three areas, i.e. economy, social and environment. On the economy, ICT will be served as an essential tool to better the living condition of citizens and eradicate poverty. On the environment, ICT will be applied to help conserve and be friendly to environment. On the social aspect, the use of ICT needs to comply with the rules and regulations and be well accepted by society and will help connect people in different regions and support borderless trading.

The essence of ICT Sustainability for Smart Thailand is the ability to utilize ICT in balance for the growth of economy, the conservation of environment and the enhancement of quality of life, social interaction, engagement and empowerment of people in society.

2.7 Proposed model for Thai HR ICT Competency Standard Framework

With the ICT competency, the Skills Framework for the Information Age (SFIA) and the Skills for the 21st Century to support the new generation of competency including the previous interviews with the focused groups of the ICT Subject Matter Experts from industry,

business, academic and government sectors from various countries in ASEAN such as Thailand, Vietnam, Malaysia and Singapore, particularly in Thailand and Vietnam which the studies has been done through the key government officials such as Ministry of ICT, Ministry of Science and Technology, Ministry of Education and Training, key ICT industry associations both local and foreign associations and major academic institutes, the model for Thai ICT Competency Standard Framework for Supporting Sustainable Smart Thailand Towards AEC 2015 could be structured as in Fig.6.

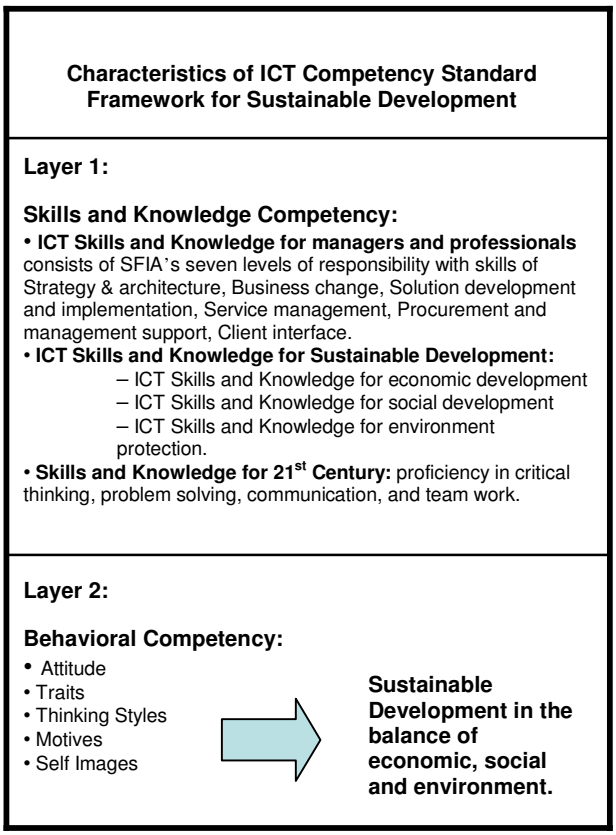


Fig.6 Components of ICT Competency

In, Layer 1, the HR ICT Competency Standard Framework for Supporting Sustainable Smart Thailand Towards AEC 2015 will consist of the ICT skills and knowledge for managers and professionals defined by SFIA, ICT skills and knowledge for sustainable development i.e. economy development, social development and environment protection and the Skills for the 21st Century.

Hence, the SFIA used as a foundation of ICT skills framework while the ICT skills (and knowledge) for sustainable development will include the ability to apply ICT to drive the economy growth – reducing and seeking to eradicate income poverty, achieving higher levels of prosperity etc.; the social development – improving the quality of education, health, housing and enhancing the quality of social interaction, engagement and empowerment etc.; the environmental protection – reducing pollution and other negative impacts on the environment, mitigating the effects of industrialization and human activity, and seeking to achieve sustainable use of resources in the interest of future generations.

However, the more important component of the competency is the one beyond Skills and Knowledge Competency which is the Behavioral Competency, i.e. Layer 2, such as Attitude, Traits, Thinking Styles and Motives towards ICT and ICT usage towards sustainable development which will be the driver to use ICT properly to achieve sustainable development with balance of economy, social and environment.

2.8 On-going studies on the Thai HRD ICT Competency Standard Framework

The on-going research on the Thai HRD ICT Competency Standard Framework for Supporting Sustainable Smart Thailand Towards AEC 2015 will be focusing on the study of gap analysis of the current competency standard of ICT skills and knowledge of ICT workforce in Thailand as compared to ones of SFIA under the survey from the targeted ICT workforce groups in each sector consisting of academic, ICT industry, business, particularly in the areas such as Logistics, Healthcare and Tourism sectors, and government sectors while the other studies from the targeted groups such as the ability and know-how to apply and use ICT to help enable the implementation of sustainable development in economy growth, social equality and environment conservation including the necessary skills for effective work in profession and carrier such as proficiency in critical thinking, problem solving, communication, and teamwork will also be done.

The study above will be implemented by the arrangement of the focused group interview through Delphi method which is designed to overcome biases from follow-the-leader tendencies and a reluctance to abandon previously stated views by allowing a group of experts to reach consensus through anonymous discussion. The study process will be done with the selected group of ICT workforce and the group of Subject Matter Experts, Policy Makers, authorities and key stakeholders from the academic, ICT industry, business and government sectors, on their views on the required ICT competencies which could be shown in a matrix of relationship between ICT competency types and job types to support the sustainable development objectives of the country.

On the second layer which is the Behavioral Competency, the study will also be done through Delphi process on the required ICT behavioral group of competencies such as attitude, self-image, traits and motives towards ICT and the application of ICT for sustainable development in their organizations and societies.

Besides the preliminary study outcomes on ICT HR competency standard conceptual framework, with the well designed data collection, interviews process, strong evaluation techniques and good statistics analysis, the research could also be represented in a mathematical model on relationship between the required ICT competencies and job type for sustainable development which could be further used and applied at the country and regional level.

The results from this study could then also be referred as a guideline for the improvement of the current the Thai ICT HRD Competency Standard and could be implemented at the policy level such as Ministry of ICT and the competency framework level such as Thailand Professional Qualifications Institute (TPQI) to assure on the successful ICT HR competency development for Smart People under the Smart Thailand policy.

In addition, for the completion of the study, the analysis on the ICT competency standard in the region and worldwide in the sustainable development area could also be used as reference on the gap bridging for the competency framework for the ICT Human Resource Development in Thailand as well.

3. Conclusions

As Thailand aims to be an integral part of the AEC 2015 as well as the potential ICT Hub and capital for ASEAN, the country has been preparing to enter the ASEAN Community with its goal to be a knowledge-based economy through the Smart Thailand 2020, particularly on the sustainable development aspect. In doing so, Thailand will need to have the right human resource with the right competencies to support the mission. The way to build up this human resource capacity for the sustainable development is to set up the framework for ICT competency standard for the workforce and study for the gap of competency between the nation's current status and the world standard one. This also includes the improvement on the competency beyond skills and knowledge such as attitude, traits, self-image and motives toward ICT and the ICT applications to support sustainable development for Thailand.

4. Suggests for further works

The research could then bring to further potential studies on the overall Thai ICT HR competencies, the strategy for ICT HRD, the ICT competencies development model, the benchmarking on ICT HR competency among ASEAN nations, the HR ICT competency development for ASEAN and other related topics concerning ICT HRD which could be used as a tool to prepare for the ICT HR to support a successful Smart People towards Smart Thailand and ASEAN Community.

References

[1] P. Trairatvorakul, "Roles and Experiences in Cultivating Global Citizens," *The International Conference on General Education*, Bangkok, Thailand, 2012.

[2] *The ASEAN Leaders adopted the ASEAN Economic Blueprint at the 13th ASEAN Summit on 20 November 2007*. Available: http://www.asean.org/archive/

5187-18.pdf [Jan, 25, 2013].
[3] United Nations Environment Program Regional Resource Centre for Asia and the Pacific, *Sustainable Development Priorities for Southeast Asia*, 2004.
[4] National Electronics and Computer Technology Center, *National ICT Policy Framework 2011-2020: ICT 2020, Ministry of ICT*, 2010.
[5] R. C. Severino, "Speeches & Statements of the Former Secretaries-General of ASEAN," *ASEAN-EU Conference sponsored by Forum Europe and the ASEAN Brussels Committee*, Brussels, 28 June, 2001.
[6] Nils-Holger Schmidt et al., "Towards a Procedural Model for Sustainable Information Systems Management, Universität Göttingen," *The 42nd International Conference of System Sciences (HICSS '09)*, Hawaii, Jan, 2009.
[7] World Commission on Environment and Development, *Our Common Future*, Oxford University Press, UK and New York, 1987.
[8] J. Elkington, Cannibals with forks: *The triple Bottom Line of 21st century business*, Captone, Oxford 1997.
[9] T. Dyllick and K. Hockerts, "Beyond the business case for corporate sustainability", *Business Strategy and the Environment*, Vol. 11, pp. 130-141, 2002.
[10] Triple bottom line, Wikipedia, Available: http://en.wikipedia.org/wiki/Triple_bottom_line, [Jan, 25, 2013].
[11] F. Spreckley, *Social Audit - A Management Tool for Co-operative Working*, 1981.
[12] Souter et. al, *ICTs, the Internet and Sustainable Development: Towards a new paradigm, International Institute for Sustainable Development (IISD)* 2010, Available: http://www.iisd.org/pdf/2010/icts_internet_sd_new_paradigm.pdf, [Jan, 25, 2013].
[13] J. Fiksel and A. Hechtouter, *Environment and Serurity: Three Pillars of Sustainable Development,* Available: http://www.eoearth.org/view/article/51cbf17c78 96bb431f6a5af0/, [May 10, 2012].
[14] United Nations General Assembly. (2005). Resolution 60/1. 2005 World Summit Outcome. Retrieved August 2010, Available: http://unpan1.un.org/intradoc/groups/ public/documents/un/unpan021752.pdf [Jan, 25, 2013].
[15] *Competency Ice Berg Model - Meaning and its Components*, Available: http://www.managementstudyguide.com/ Competency -iceberg-model.htm [Jan, 25, 2013].
[16] *The definition and selection of key competencies*, *Executive Summary*, OECD, 2005, Available: http://www.oecd.org/ edu/statistics/deseco [Jan, 25, 2013].
[17] *What is IT Competency*, Available:http://www.wisegeek.com/ what-is-it-competency.htm [Jan, 25, 2013].
[18] Nuangjamnong et al., "The Possibilities of an Implementation National ICT Competency Framework for Thailand ICT Master Plan," *The 2nd international Conference on Education Technology and Computer (ICETC)*, 2010.
[19] Civil Service, *Competency and skills framework*, Available: http://www.civilservice.gov.uk/ my-civilservice/networkslprofessional/itl [Jan, 25, 2013].
[20] Wikipedia, *Skills Framework for the Information Age*, Available: http://en.wikipedia.org/wiki/ Skills_Framework_for_the_Information_Age [Jan, 25, 2013].
[21] SFIA Foundation, *SFIA 5 framework reference Skills defined in categories and subcategories*, 2011.
[22] Ledward, B. C., and D. Hirata. *An overview of 21st century skills*, *Summary of 21st Century Skills for Students and Teachers*, by Pacific Policy Research Center. Honolulu: Kamehameha Schools–Research & Evaluation, 2011.
[23]ICT Sustainability for Smart Thailand, "Speech of Minister of ICT of Thailand," *CIO Conference*, Bangkok, Thailand, November, 2012.

A Comparison of the Need and Readiness of Laptops and Tablets Usage for Working at Tesco Lotus and BAAC Bank in Thailand

Sanon Chimmanee † **Sunya Chimmanee †**

†M.S. ITM-ONLINE, Information Technology Faculty Rangsit University,
sanon.s@rsu.ac.th, sunya_kimeng@hotmail.com

Abstract

Nowadays, tablets are widely used in Thailand. Hence, most organizations or companies currently need to know whether they could replace laptop computers in any workplace or not. Thus, the purposes of this paper are as follows: 1) to compare the needs in Laptops and Tablet usage in different population clusters. 2) To investigate the readiness of using Tablets and functional capability of Tablet for working. First population cluster is Tesco Lotus employees at a head quarter in Thailand. The results found that in average, 93% of Tesco Lotus employees is preferred the Laptops to Tablets. A mean of ability of the usage for working is 4.38 while a mean of the Tablet usage is 3.04. The other population cluster is the BAAC Bank employees at a head quarter. The results found that in average, 89% of Tesco Lotus employees is preferred the Laptops to Tablets. A mean of ability of the Laptops usage for working is 4.37 while a mean of the Tablet usage is 3.10. It is interesting to note that the both government organization and public company in this paper have the same direction.

Keywords: Tablets, Laptop, Tesco Lotus, BAAC bank

1 Introduction

At present, Tablets are become increasingly popular for both organization and company. United and Continental Airlines were discarded their paper flight manuals and navigational charts and embracing the iPad. Approximately 11,000 iPads will be given to all United and Continental pilots starting on August 2011[1]. "As the result of all staff having a tablet, they are now utilizing an Internet-based calendar for scheduling of all meetings and appointments" said by Greg Crowe [2]. Worker can work on their tablets while bidding time at court or in a waiting room while a child is visiting a counselor or doctor [2]. Michael Endler mentioned that "We're still learning a lot about tablets -- what benefits they bring to business, where they don't work and what their limitations are? " [3]. Mark Kyrnin stated that "In many ways, they can almost replace a laptop for someone on the go. But is a tablet really a better choice for someone over a more traditional laptop?"[4].

Recently, Thai government has launched the pilot project in 2012, which gives a free tablet for a 1st grade student. Achieving the readiness of using tablet in teachers and parents who are in charge of management for supporting students' education is very important. Thus, the readiness examination should be identified. In [5], Meennapa Rukhiran and Sanon Chimmanee focused on investigating readiness of teachers and parents of grade 1st students that were examined from four districts under administrative divisions; Mueang Chathaburi, Tha Mai, Na Yai Am and Kaeng Hang Maeo in Chanthaburi Province during October to November 2012 [5]. Interestingly, the teachers have the readiness in Tablets usage for supporting children in a moderate level with MEAN 3.20 while parents have the readiness in a high level with MEAN 3.56. Peerada Yuyuenyong and Sanon Chimmanee studied the comparison between usages of the PC computer/laptop and the tablet for officers of the Thai E-Customs department [6]. The result found that the usage of the PC computer/laptop is preferred than the tablet in the function of importing and exporting up to 11 % because the PC computer/laptops is more effective than tablet e.g., typing, memory, displaying.

What Factors Contributed to the Success of Apple's iPhone?[7]. John Laugesen and Yufei Yuan defined a set of success criteria to observe the success of the iPhone and present a comprehensive success model [7]. Jutarak Kunasakpanich studied factors affecting the selection of

the smart phone between iPhone and Backberry of Rangsit University students during August to September 2012 [8]. There were 400 samples from total 25,424 students. The result found that iPhone is popular than BlackBerry. For iPhone, the most influential factor for iPhone selection is a brand itself. In [9], Saranya Ruensri studied the behavior choice to buy mobile phone between blackberry and other mobile phones of students in Chiang Mai University. The information was of primary source coming from questionnaire interview of 400 samples from total 24,592 students during May 2010. It was found that functional capacity of the product factor is the most influential factor.

This paper investigates to know whether organizations or companies could replace laptop computers in any workplace or not. Thus, the purposes of this paper can be stated as follows: 1) to compare the needs in using laptops and tablets in different population samples 2) to examine the readiness of the use of tablets as well as their functional working capabilities. First population cluster is Tesco Lotus employees at a head quarter in Thailand. There were 332 samples from total 1,804 employees. These questionnaires were done during November to December 2012. The results found that in average, 93% of Telco Lotus employees preferred Laptops to Tablets. A mean of ability of the Laptops usage for working is 4.38 (high level) while the mean of the Tablets usage is 3.04 (moderate level).
The other population cluster is the Bank for Agriculture and Agricultural Co-operatives (BAAC). There were 175 samples from total 220 staffs at IT department. These questionnaires were done during November 2012 to January 2013. The results found that in average, 89% of Telco Lotus employees also preferred Laptops to Tablets. A mean of ability of the Laptops usage for working is 4.37 (high level) while the mean of the Tablet usage is 3.10 (moderate level).

It is interesting to note that the both government organization and public company in this paper are the same direction. Apparently, all functional working capabilities of laptops appear to be higher than those of tablets except for the social network purposes.

The rest of this paper is organized as follows. Section 2 presents literature reviews. Research methodology is stated in the section 3. Experimental result is presented in the section 4. Section 5 gives conclusion and discussion to the paper.

2 Literature Reviews

2.1 Tablet

United and Continental Airlines were discarded their paper flight manuals and navigational charts and embracing the iPad. Approximately 11,000 iPads will be given to all United and Continental pilots starting on August 2011[1]. Recently, the Thai government has launched the pilot project in 2010, which gives a free tablet for a 1st grade student. Achieving the readiness of using tablet in teachers and parents who are in charge of management for supporting students' education is very important.

2.2 Readiness in Tablets Usage in Thailand

Meennapa Rukhiran and Sanon Chimmanee focused on investigating readiness of teachers and parents of grade 1st students during October to November 2012. There are 88 schools with 110 teachers and 2,297 students that are examined from four districts under administrative divisions; Mueang Chathaburi, Tha Mai, Na Yai Am and Kaeng Hang Maeo in Chanthaburi Province. There were 110 teachers and 341 parents who were the samples [5]. The result found that the teachers have the readiness in Tablets usage for supporting children in the moderate level with MEAN 3.20 while parents have the readiness in the high level with MEAN 3.56.

In [6], Peerada Yuyuenyong and Sanon Chimamanee studied of abilities, expectations and attitudes of applying of the tablet in the E-Customs i.e., paperless system in 2012. A study of comparison between usage of the computer/laptop and the tablet for officer of the E-Customs department was also investigated. The simple size of 203 was drawn from 397 officers of the Laem Chabang Port Customs of officer in Chonburi province. The results show that 1) approximately 65 % of customs officer can use the web and e-mail service via the tablet to perform the functional import and export. 2) The expected and attitudes of the tablet usage in E-Customs is high about 70% for helping them in working with E-Customs. 3) The usage of the computer/laptops is preferred than the tablet in the function of importing and exporting up to 11 % because the computer/laptop is more effective than tablet e.g., typing, memory, displaying.

2.2 Influential Factor for Product Selection

In [8], Jutarak Kunasakpanich studied factors affecting the selection of the smart phone between iPhone and Backberry of Rangsit University students during August to September 2012. Out of 25, 424 students, only 400 students were used as the subjects for this study. The result found that iPhone is popular than BlackBerry. For Blackberry, the most influential factor is the price, the distribution channel is second high and the social factor is the lowest. For iPhone, the most influential factor for iPhone selection is a brand itself. The product factor is second high and the social factor is also the lowest.

In [9], Saranya Ruensri studied the behavior choice to buy mobile phone between blackberry and other mobile phones of students in Chiang Mai University. The information was of primary source coming from questionnaire interview of 400 samples from total 24,592 students during May 2010. It was found that product, price, buying objective, functional capability of the product, and the current popularity of specific mobile phone models were associated with the decision to buy Blackberry or mobile phone. Functional capacity of the product factor is the most influential factor.

3 Research Methodology

In this paper, there are four main objectives as follow:

- To observe the needs of laptops and tablets usage.
- To investigate the readiness of using laptops and tablets.
- To examine functional working capabilities of laptops and tablets.
- To compare the objectives of 1, 2, and 3 between Tesco Lotus employees at a head quarter in Thailand that is a public company and Bank for Agriculture and Agricultural Co-operatives (BAAC) that is Thai government bank.

3.1 Population

There are two different population clusters. First population cluster is Tesco Lotus employees at the head quarter in Thailand. Out of 1,804 employees from 6 departments, only 332 employees were used as the subjects for this study. The other population cluster is 175 samples from 220 staffs in IT department of BAAC bank.

3.2 Conceptual Framework

Independent variables:
Sample population:

- Tesco Lotus employees at a head quarter in Thailand
- Officers of BAAC bank

Devices:

- Laptops
- Tablets

Dependent variables:

- Need in Tablets or Notebook
- Readiness of Laptops and Tablets usage
- Functional Capability of Tablets and Laptops for working

3.3 Tool design

Because functional working of Tesco Lotus and BAAC bank is different, this paper designs specific questionnaires for each place. Both questionnaires contain 4 parts as follow:

- Personal information
- Need in laptops or tablets
- Readiness of laptops and tablets usage
- Functional working capability of laptops and tablets

3.4 Collect Data

For Telco Lotus, these questionnaires were done during November to December 2012. For the Bank for Agriculture and Agricultural Co-operatives (BAAC), these questionnaires were done during November 2012 to January 2013.

3.5 Analysis Data

This paper uses statistic mathematic that are frequency, percentage, mean, and standard deviation.

4 Experimental Result

There are six subsections: subsections 4.1 to 4.3 are about Telco Lotus. And subsections 4.4 to 4.6 are about BAAC bank.

4.1 Tools Selection for Telco Lotus

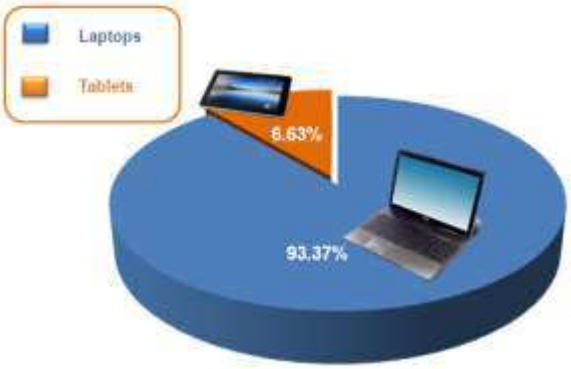


Figure 1. Tools Selection for Telco Lotus

In Figure 1, it shows that in average, 93% of Telco Lotus employees preferred laptops to tablets.

4.2 Comparison of Readiness of Usage for Telco Lotus

Table1. Comparison of readiness in usage between Laptops and Tablets for Telco Lotus

Readiness of Usage		Laptop		Tablet	
		Mean	S.D	Mean	S.D
1	Hardware and options	4.01	0.75	3.75	1.00
2	Basic Applications	3.37	0.71	4.17	0.70
3	Application Installations	3.71	1.22	2.95	1.32
4	Troubleshooting	4.23	0.71	2.90	0.77
5	Maintenance	3.37	1.12	2.63	0.86
6	General usage	4.06	0.76	2.85	0.96
Total		3.80	0.96	3.21	1.11

From Table 1, it is seen that all features of the readiness in usage of Laptops are higher than Tablets except the basic application feature.

4.3 Comparison of Functional Working Capability for Telco Lotus

From Table 2, it is seen that all functional working capabilities of laptops are higher than tablets except the social network feature.

Table2. Comparison of Functional capability between Laptops and Tablet for Telco Lotus

Functional working capabilities		Laptop		Tablet	
		Mean	S.D	Mean	S.D
1	Internet	4.63	0.55	4.40	0.75
2	Email	4.66	0.64	4.25	0.93
3	MS Word	4.60	0.52	2.65	0.70
4	MS Excel	4.59	0.53	2.58	0.81
5	MS Power-Point	4.58	0.61	2.46	0.81
6	Sale System (GMIS)	3.94	1.08	2.28	1.16
7	General System (RAMS)	4.18	0.85	2.19	1.08
8	File Sharing	4.18	0.92	2.42	0.93
9	Typing	4.42	0.73	2.61	1.01
10	Social Network	4.02	1.09	4.61	0.70
Total		4.38	0.82	3.04	1.28

4.4 Tools Selection for BAAC bank

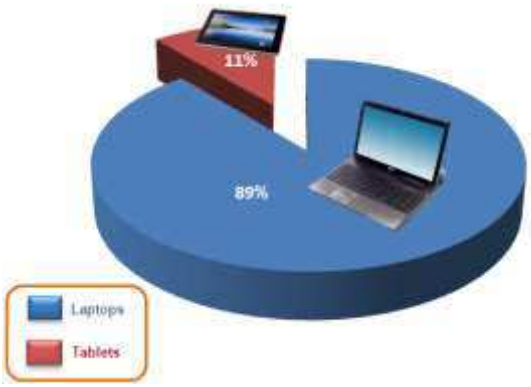


Figure 2. Tools Selection for BAAC bank.

In Figure 2, it displays that in average, 89% of BAAC staffs preferred laptops to tablets.

4.5 Comparison of Functional working Capability for BAAC Bank

Table 3. Comparison of Functional working capability between laptops and tablets for BAAC bank.

Functional working capabilities		Laptop		Tablet	
		Mean	S.D	Mean	S.D
1	CPU process	4.20	0.83	3.45	1.08
2	Memory	4.14	0.84	3.26	1.08
3	Option Connection	4.24	0.87	2.46	1.28
4	OS	4.34	0.77	3.16	1.11
5	Microsoft Office	4.51	0.77	2.50	1.21
6	BAAC Applications	4.34	0.89	2.24	1.24
7	BAAC Intranet	4.56	0.74	3.21	1.28
8	BAAC Email	4.53	0.73	3.55	1.23
9	Remote Access	4.24	1.04	2.12	1.32
11	Electronic Document Access	4.37	0.89	2.67	1.21
12	Websites	4.53	0.75	4.20	1.04
13	Social Network	4.40	0.90	4.28	1.00
Total		4.37	0.83	3.11	1.15

From Table 3, it is shown that all functional working capabilities of laptops are higher than tablets.

4.6 Comparison of Readiness in Usage for BAAC Bank

Table 4. Comparison of readiness of usage between laptops and tablets for BAAC bank.

Readiness of Usage		Laptop		Tablet	
		Mean	S.D	Mean	S.D
1	Hardware and options	4.11	0.73	3.51	1.04
2	Operating System (OS)	4.11	0.80	3.52	1.07
3	File Type	4.42	0.71	3.85	1.11
4	Email	4.42	0.66	4.17	0.95
5	Recycle bin	4.28	0.85	3.73	1.09
6	Basic Application	4.48	0.67	4.05	1.05
7	Application Installation	4.22	0.93	3.75	1.13
8	Troubleshooting	4.20	0.92	3.60	1.13
9	Maintenance	4.28	0.79	3.84	1.06
Total		4.28	0.78	3.78	1.07

From Table 4, it can be seen that the readiness of laptops usage on all features appear to be higher than tablets.

5 Conclusion and Discussion

5.1 Conclusion

For Telco Lotus, the results found that in average, 93 % of employees preferred Laptops to Tablets. A mean of ability of the Laptops usage for working is 4.38 (high level) while a mean of the Tablet usage is 3.04 (moderate level). For BAAC bank, the results found that in average, 89% of officers also preferred Laptops to Tablets. A mean of ability of the Laptops usage for working is 4.37 (high level) while a mean of the Tablet usage is 3.10 (moderate level). More interestingly, all features of MEAN functional working capabilities of Laptops are higher than Tablets except the social network feature for Telco Lotus. For BAAC bank, MEAN social network of Laptops and Tablets are nearly the same. This means that the laptop is still powerful for working and Tablets still may be an option if you use it mainly for entertainment, social network and web usage. The size of population should be extended to other clusters in order to get more information in the further work.

5.2 Discussion

In [2], Greg Crowe stated that the tablets, which helped the workers to be able to enter and update any information from the distance, allowed them to have more time with their families. Overall, workers have indicated a time savings of anywhere from an hour a day to feeling as if they gain a day and a half to two additional days a week in time saved [2]. The results of this paper are seen that Tablet is appropriate to fast connection, and simply application especially in social network. However, it may be not convenience to complex applications for working. It is the same direction as Mark Kyrnin [4]. He stated that “laptops can be extremely portable and have a much wider range of tasks they can be used for”. Additionally, most people cannot type as quickly or as accurately on a virtual keyboard. There is no an option of adding an external Bluetooth keyboard to most tablets to make this more like a laptop [4],[6].

Acknowledgment

Authors many thank to Mr.Watchara Jamjumras for collecting information from Tesco lotus employees at Headquarter in Thailand. Authors also a big thank to Mr.Nut Pispratueng for collecting information from officers of BAAC bank.

References

[1] Kelly Hodgkins, United and Continental Airlines to hand out 11,000 iPads to its pilots. http://www.tuaw.com/2011/08/23/united-and-continen-tal-airlines-to-hand-out-11-000-ipads-to-its/, Aug 23rd 2011, July, 2013

[2] Greg Crowe, Tablet suite saves case workers their most valuable commodity, GCN Technology, Tools and Tactics for Public Sector IT, Apr 16, 2013 http://gcn.com/Articles/2013/04/16/Tablet-suite-saves-case-workers-time.aspx?Page=2, July, 2013

[3] Michael Endler, The Good And Bad Of Tablets At Work, InformationWeek.com http://www.informationweek.com/mobility/business/the-good-and-bad-of-tablets-at-work/240156604, July, 2013.

[4] Mark Kyrnin, Tablet PCs vs. Laptops Should You Buy a Tablet PC Instead Of a Laptop, PC Hardware-Reviews, about.com, http://compreviews.about.com/od/buyers/a/Tablets-vs-Laptops.htm, July 2013.

[5] Menapa and Sanon Chimmanee, Teachers and parents' readiness for supporting tablet use in 1st grade classroom of Chanthaburi Primary Educational Service Area Office 1, Thailand, BUU international conference 2013 at Jontiem Plam Beach, Pattaya, 4-5 July, 2013, Thailand.

[6] Perada and Sanon Chimmanee, Ready in the application of the Tablet in the Thailand E-CUSTOMS, The Fourth National Graduate Research Conference (NGRC), Ramkhamhaeng University, 30 August, 2012, Thailand.

[7] John Laugesen, and Yufei Yuan, What Factors Contributed to the Success of Apple's iPhone, IEEE Computer Society, Proceeding of ICMB-GMR '10 Proceedings of the 2010 Ninth International Conference on Mobile Business / 2010 Ninth Global Mobility Roundtable, Pages 91-99, 13-15 June 2010, Athens

[8] Juthamak Kunasakpanich, A study of behavior and factors affecting the selection of the smart phone between iPhone and Backberry of Rangsit University students, Independent Study (IS), M.S.ITM-Online, Rangsit Unversity in 2012.

[9] Saranya Ruensri, The comparison of behavior choice to buy mobile phone between BlackBerry and other mobile phones of students in Chiang Mai University, Independent Study (IS), Master of Economics, Chiang Mai University in 2010.

A Case Study: Thailand Customs Electronic System (TCES) In Aranyaprathet Customs House

Savika Nak-on **Monruthai Radeerom**

Faculty of Information Technology,
Rangsit University, Pathumthani
12000, Thailand
wift3@hotmail.com
m.radeerom@rsu.ac.th

Abstract

The Thailand Customs Electronic System (TCES), formerly known as "e-Customs", has been used since 2007. In order to gauge the effectiveness of the system, data was collected between August 2012 to June 2013 from 40 customs officers and 15 customs brokers through simple random sampling. They were asked to answer questions using a rating scale assessing the performance of TCES. That data was analyzed using descriptive statistics by arithmetic mean, and showed the following: (1) an average of 4.13 (high level) and S.D. 0.60, (2) those persons sampled are knowledgeable about the system at a "moderate" level yet show expertise for the job they are assigned, and (3) that age, responsibility (custom officers), category, duties of the position, import – export form (Customs Broker), etc. do not affect the system's performance.

Keywords: effectiveness; custom electronic system; quantitative research

1 Introduction

The world has entered an era of globalization and rapid change perfect for the advancement of technology and innovation. The current system of global trade relations has linked the domestic economy and the global economy. Organizations like the Free Trade Area (FTA) have intensified economic competition and are now major factors linking the economies of most countries. Thai Customs has implemented TCES to facilitate international trade and it is an important tool for processing international shipments smoothly, conveniently, and quickly. Therefore, it is crucial to study the opinions and feedback of the employees which run the TCES, in order to correct problems and implement more efficient solutions in a timely manner.

Recently, Peerada Yuyuenyong and Sanon Chimmanee studied the comparison between usages of the PC computer/laptop and the tablet for officers of the Thai E- Customs department [1]. The result found that the usage of the PC computer/laptop is preferred than the tablet in the function of importing and exporting up to 11 % because the PC computer/laptops is more effective than tablet e.g., typing, memory, displaying.

In [2], Manoch Rodsom studied of Paperless system for Thai Customs and results in more efficient, responsive to customer's needs, and to establish a transparent system of government.

Walaiporn Rukkhachart [3] observed an implementation of Paperless System in Public Agency at The National Health Security Office [3]. It is found that the response of those staff members has been extremely positive according to surveys.

This paper investigates opinion on the efficiency of TCES Aranyaprathet Sakaeo province from Custom officers and brokers.

From experimental result of Custom officers, it can be seen that the efficient international standard is the highest while the efficient in reducing the operating costs of enterprises is the lowest. From experimental result of Custom brokers, it is found that the efficient transparency is the highest while the efficient of facilitation of international trade is the lowest.

The rest of this paper is organized as follows. Section 2 presents literature reviews. Research methodology is stated in the section 3. Experimental result is presented in the section 4. Section 5 gives conclusion to the paper.

2 Literature Review

2.1 Thai Government Has Focused on the Use of Information Technology.

Manoch Rodsom stated that the policy of reforming the Thai government has focused on the use of information technology to provide accurate input for making tangible improvements quickly. Government agencies must expedite research in the area of information technology to reform management systems to achieve efficiency and effectiveness. Customs is one such government agency that advances information technology applications in the development and reformation of management systems. A perfect example of these major reforms was the implementation of the "Electronic Data Interchange (EDI)", which has resulted in improved times for customs clearance by eliminating unneeded procedures and created greater transparency. According to the Rodsom study, Customs developed EDI to be paperless, more efficient, responsive to customer's needs, and to establish a transparent system of government.

2.2 Paperless Systems within the Process.

In 2003, Walaiporn Sukchat found that using paperless systems reduces processing time and the workload of the staff. In addition, it vastly improved internal communication, plus reduced the cost of printing, copying, and the working hours of staff. The response of those staff members has been extremely positive according to surveys.

On February 13, 2012 Customs Department modified the e-Customs system to be the Thailand Customs Electronic System (TCES) in order to integrate all of the electronic systems of the Customs Department. The goal was to prepare for liberalization of trade within the ASEAN Community and facilitate international trade benefiting entrepreneurs with competitiveness, low cost, convenience, speed, accurate performance, and efficiency.

2.3 Thailand E-Customs and Tablet

In [4], Peerada Yuyuenyong and Sanon Chimamanee studied of abilities, expectations and attitudes of applying of the tablet in the E-Customs i.e., paperless system in 2012. A study of comparison between usage of the computer/laptop and the tablet for officer of the E-Customs department was also investigated. The simple size of 203 was drawn from 397 officers of the Laem Chabang Port Customs of officer in Chonburi province. The results show that 1) approximately 65 % of customs officer can use the web and e-mail service via the tablet to perform the functional import and export. 2) The expected and attitudes of the tablet usage in E-Customs is high about 70% for helping them in working with E-Customs. 3) The usage of the computer/laptops is preferred than the tablet in the function of importing and exporting up to 11 % because the computer/laptop is more effective than tablet e.g., typing, memory, displaying.

2.4 Effects of e-Customs

In [5], Panida Peetiboonpaiboon found that the effects of the use of e-Customs are as follows: 1) amount of time spent in dealing with exports and import was lessened. 2) Paper used for documents decreased by 50%. 3) Information provide can now be verified even after goods have been released. In conclusion, the use of e-Customs in customs procedures facilitates trading, aids in economic development, and accommodates exporters, thereby contributing to the expansion of international trade.

In [6], Hataiporn Inta Indicated that functionality, usability, efficiency, maintainability and portability were at a high level. Product operation, product revision and product transition also were rated highly. 23 criteria such as assess auditing, access control, communication commonality etc. were all rated highly. Opinion on the level of importance the key success factors for implementing the e-Customs system in finance and accounting were excusive policy, software, manual user, technology, people, internal control and data. Few problems were found for implementing the e-Customs system in finance and accounting.

3 Research Methodology

3.1 Conceptual framework

To study the efficiency of TCES within two groups: a) the customs officers have the factors 1) responsibility 2) age, and b) the customs brokers 1) category 2) position 3) the average

monthly entry clearance and 4) the average monthly export clearance. Review of literature can establish the following concepts (see figures 1).

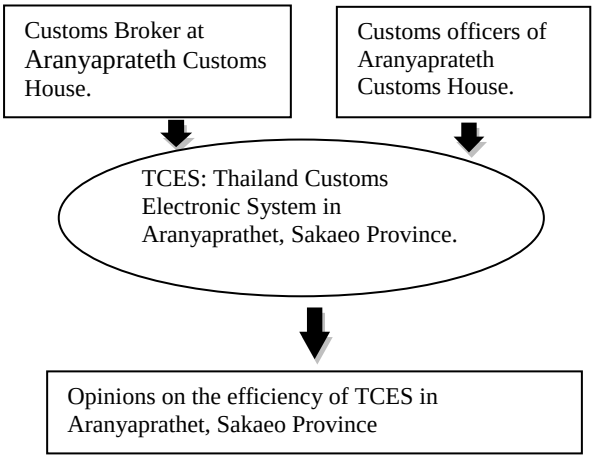


Figure 1. Conceptual framework of a figure.

3.2 Hypotheses

It is summarized in 5 hypotheses (see Table 1).

Table 1: Research Hypotheses

No.	Notes
1.	The facilitation of international trade is efficient.
2.	Reducing the operating costs of enterprises is efficient.
3.	The international standard is efficient.
4.	Increasing trade competitiveness is efficient.
5.	Transparency is efficient.

3.3 Methodology

Quantitative research (questionnaire) as well as qualitative (interview) was used to study the opinions and problems in depth.

3.4 The questionnaire

The questionnaire about the efficiency of TCES is tested by reliability (α) 0.85 of Cronbach's Alpha Coefficient. There are 3 parts as

- General Information of samples consists of age, responsibility, etc.
- Opinions from open-ended questions that have rating scale by Likert for 5 ranking therefore as 4.51 – 5.00 = Highest, 3.51 – 4.00 = High, 2.51 – 3.00 = Moderate, 1.51 – 2.00 = Low and 1.00 – 1.50 = Lowest.

3.5 Data Analysis

Data analysis used descriptive statistics about the effectiveness of the TCES by arithmetic mean ( ).
After collecting all information, our evaluating method used the one way analysis in term of variance (s2) = 0.05 and compared their results with least significant difference (LSD).

3.6 Selecting Samples for Research

This study used survey methodology (questionnaire) for 2 groups of random samples.
First group is 40 samples from a population of 60 customs officers at Aranyaprathet Customs House.
Second group is 15 samples from a population of 150 people who are serviced by officers. In addition, a structural insights pattern (interview) was used.
Two groups were separated for the interview. One is 5 samples from the 40 customs officers at Aranyaprathet Customs House and the other is 10 samples from the 15 Customs Brokers.

4 Experiment Result

Sample of 55 participants at Aranyaprathet Customs House are in their 50s (45% majority age than 50 years). They are mostly responsible for good inspection (65%). 60 percent of the customs Broker are agents. 73% of the customer files the entry declaration for about 1-50 copies, whereas 47.60% files the exit declaration about 151 copies on average.
Research results of the second part in the questionnaire (see Table 2).

Table 2: Opinion on the efficiency of TCES in Aranyaprathet Sakaeo Province from Customs Officer

No.	Table Column Head	$\bar{x}$	***S.D.***	***Level***
1.	The facilitation of international trade is efficient.	4.06	0.59	High
2.	Reducing the operating costs of	3.85	0.73	High

No.	Table Column Head	$\bar{x}$	*S.D.*	*Level*
	enterprises is efficient.			
3.	The international standard is efficient.	4.19	0.54	High
4.	Increasing trade competitiveness is efficient.	4.13	0.61	High
5.	Transparency is efficient.	4.05	0.53	High
	Total	4.05	0.6	High

From table 2. Custom officers' opinions about efficiency in 5 aspects of TCES are in high level in average of 4.06 and 0.596 S.D. level (4.19), followed by trade competitiveness (4.13), the facilitation of international trade (4.06), transparency (4.05) and reducing the operating costs of enterprises(3.85), respectively.

Most of the staff members interviewed showed consistency with the data obtained from questionnaires. It appears TCES is very effective in all 5 aspects, but the linkage between government agencies to approve the import – export goods in the form of Single Window (One Stop Service) is not perfect, since some products and agencies are still not connected to the system. The system is still evolving, and more agencies are being integrated.

In addition, most officials only have a moderate understanding of TCES due to their limited knowledge of the system as a whole. The Customs Department should organize a seminar concerning the theory and practices of TCES to enhance knowledge and skills of the officials and staff.

Research results of the third part in the questionnaire (see Table 3).

Table 3: Opinion on the efficiency of TCES in Aranyaprathet Sakaeo Province from Customs Broker

No.	Table Column Head	$\bar{x}$	*S.D.*	*Level*
1.	The facilitation of international trade is efficient.	4.12	0.62	High
2.	Reducing the operating costs of enterprises is efficient.	4.17	0.67	High
3.	The international standard is efficient.	4.20	0.56	High
4.	Increasing trade competitiveness is efficient.	4.27	0.61	High
5.	Transparency is efficient.	4.29	0.53	High
	Total	4.21	0.59	High

From table 3. Customs Brokers' opinions average 4.21. After considering each aspect we found that transparency is highest average level (4.29), followed by trade competitiveness (4.27), reducing the operating costs of enterprises (4.17), and the facilitation of international trade (4.12), respectively.

Most of the customs brokers' interviews are consistent with the data obtained from questionnaires. TCES is very effective in all 5 aspects, but the linkage between the government, bank, and entrepreneur to approve the import – export products through One Stop Service (both facilitation of international trade and international standard) is not perfect, and efficiency suffers because some the government agencies and banks can connect to the system. Many are still connecting to the system and legislation is being updated to support applications, so officers still use paper documents for customs clearance and time is wasted filing these Customs entries. Most brokers have a better understanding about TCES, but this comprehension is still just moderate despite the private sector organizing the seminar about TCES. The more people involved in the process that attend these seminars, the better the understanding will be.

In addition, the various factors related to the samples do not differentiate opinions about the efficiency of TCES.

In this study, it was used as the guidelines to help improve and develop Customs system, especially TCES to be modern, convenience and efficiency as well. Moreover, it has completed the goal of Customs to increase its competitiveness.

5 Conclusion

Although customs officers and customs brokers at Aranyaprathet Customs House have positive opinions about the performance of TCES, problems still exist. These include lack of knowledge about information systems, the system still not being completed, and issues with integration, amongst others. Customs and other relevant departments should solve these problems in a timely manner to prevent future problems. They should also resolve the inadequate linkage of information between government agencies, and adjust the relevant laws to provide consistency in facilitating customs clearance. Finally, more training and seminars providing more information to those involved with TCES will enhance their operational ability even more.

References

[1] Customs Department. Thailand Customs Electronic System (TCES). *Customs Bulletin*, 2012.

[2] ManochRodsom. Governmental Reform by Information Technology: A Case Study: Paperless Customs System. Master of Arts (Political Science), *Chulalongkorn University*, 2000.

[3] Walaiporn Rukkhachart. The Implementation of Paperless System in Public Agency: A Case Study: The National Health Security Office. Master of Arts (Political Science), *Chulalongkorn University*, 2003.

[4] Perada Yuyuenyong and Sanon Chimmanee. Ready in the application of the Tablet in the Thailand E-CUSTOMS, The Fourth National Graduate Research Conference (NGRC). *Ramkhamhaeng University*, 2012.

[5] Panida Peetiboonpaiboon. The Effects of the Use of e-Customs. Master of Economics, *Ramkhamhaeng University*, 2009.

[6] Hataiporn Inta. Problem and Key Success Factors of Implementing e-Customs System in Finance Accounting in Customs House of Customs Reginal Bureau. Master of Accounting, *Chaingmai University*, 2011.

The Study of Need and Factors Affecting the Selection of Tablet between IPAD and Samsung Galaxy Tab compared between the Flight of Army Aviation Center's Pilots and Rangsit University's students

Pattama Noumthanong† Sanon Chimmanee†Sunya Chimmanee †

†M.S. ITM-ONLINE, Information Technology Faculty
Rangsit University,
pattama.n@rsu.ac.th, sanon.s@rsu.ac.th, sunya_kimeng@hotmail.com

Abstract

This paper aims to investigate the need in tablet usage between iPad and Samsung Galaxy Tab in different population clusters. One population contains students of Rangsit University while the other population consists of the Army Aviation Center's pilots. Various factors affecting the selection of tablet between two popular tablets for two population clusters were observed. The results found that the average of students of Rangsit University is preferred iPad than Samsung Galaxy Tab. On the other hand, the average of pilots of Army Aviation Center's Pilot is preferred Samsung Galaxy Tab. One major finding is that the main purpose of tablet usage for the students is information searching while that for the pilots is reading and writing information.

Keywords: Tablets, The new iPad, Samsung Galaxy Tab 7.7, Factors affecting, Pilot

1 Introduction

Nowadays, tablets are widely used for both students and workers. There is a needed to study factors affecting selection of tablets usage. The priority of various factors affecting the use of tablets is also interesting. iPad by Apple and Samsung Galaxy Tab by Samsung are very popular tablet. United and Continental Airlines were discarded their paper flight manuals and navigational charts and embracing the iPad. Approximately 11,000 iPads will be given to all United and Continental pilots starting on August 2011[1]. Recently, the Thai government has launched the pilot project, which gives a free tablet for a 1st grade student. Achieving the readiness of using tablet in teachers and parents who are in charge of management for supporting students' education is very important. Thus, the readiness examination was identified in [2].

Meennapa Rukhiran and Sanon Chimmanee focused on investigating readiness of teachers and parents of grade 1st students during October to November 2012. There are 88 schools with 110 teachers and 2,297 students that are examined from four districts under administrative divisions; Mueang Chathaburi, Tha Mai, Na Yai Am and Kaeng Hang Maeo in Chanthaburi Province. There were 110 teachers and 341 parents who were the samples [2].

In [3], JuthamakKunasakpanich studied factors affecting the selection of the smart phone between iPhone and Backberry of Rangsit University's students during August to September 2012.There were 400 sample students from total 25,424 students. The result found that iPhone is popular than BlackBerry. For iPhone, the most influential factor for iPhone selection is a brand itself. In[4], Saranya Ruensri studied the behavior choice to buy mobile phone between blackberry and other mobile phones of students in Chiang Mai University. The information was of primary source coming from questionnaire interview of 400 samples from total 24,592 students during May 2010. It was found that functional capacity of the product factor is the most influential factor.

What factors contributed to the success of iPad and Samsung Galaxy Tab. Thus, in this paper investigate the need in tablet usage between iPad and Samsung Galaxy Tab in two population clusters. Various factors affecting the selection of tablet are presented. The priority of various factors affecting the use of tablets is also investigated.

First population cluster is students of Rangsit University.There were 400 sample students from total 25,424 students. These questionnaires were done duringSeptember 2012 to January 2013.The

results found that in average, 78% of Rangsit University's students preferred iPad to Samsung Galaxy. Interestingly, the purpose of using tablet is mainly for information searching which can be accounted for 56.75%.The most influential factor for iPhone selection is a brand itself and the MEAN equals to 4.45.

The other population cluster is the Army Aviation Center's Pilot.There were 175 pilot samples. This part was done during September 2012 toJanuary 2013. From the result, it is found that the functional capacity of the product factor is the most influential factor. Interestingly, the target of using tablet is mainly for reading and writing information which can be accounted for 52 %. It is interesting to note that the different population clusters are different the purpose of using tablet and the requirement for tablet selection.

The rest of this paper is organized as follows. Section 2 presents literature reviews. Research methodology is stated in the section 3. Experimental result is presented in the section 4. Section 5 gives conclusions and discussions to the paper.

2 Literature Reviews

2.1 Tablet

In [5], Micah Stickel used the tablet PC as a lecturing device in 2009. Since the general student body contains many different types of learning styles, these new tools can provide a means to reach a wider range of students than the traditional blackboard-based lectures. Overall, the students in all courses (a total sample size of 280) were quite positive about the use of the tablet, with 56% of the students stating that "The technology of the tablet…greatly improves my learning experience." It was found that there were no statistically significant differences in how the students of different learning styles responded to the use of the tablet PC.Recently, the Thai government has launched the pilot project, which gives a free tablet for a 1st grade student. Achieving the readiness of using tablet in teachers and parents who are in charge of management for supporting students' education is very important. Thus, the readiness examination was identified in [2]. Meennapa Rukhiran and Sanon Chimmanee focused on investigating readiness of teachers and parents of grade 1st students during October to November 2012. There are 88 schools with 110 teachers and 2,297 students that are examined from four districts under administrative divisions; Mueang Chathaburi, Tha Mai, Na Yai Am and Kaeng Hang Maeo in Chanthaburi Province. There were 110 teachers and 341 parents who were the samples [2].More interestingly, from this study many teachers and parents are potentially worries regarding the limited capacity of young students' caring tablet. For example, screen and equipments caring, tablet cleaning, tablet losing and tablet crashing. In addition, managing time of using tablet should be controlled by teachers and parents because students can suffer eye injuries from spending time too long on tablet.

2.2 Influential Factor for Product Selection

What Factors Contributed to the Success of Apple's iPhone?[3]. John Laugesen and Yufei Yuan defined a set of success criteria to observe the success of the iPhone and present a comprehensive success model [3].

In [4], Jutharat Kunasakpanich studied factors affecting the selection of the smart phone between iPhone and Backberry of Rangsit University's students during August to September 2012. Out of 25, 424 students, only 400 students were used as the subjects for this study. The result found that iPhone is popular than BlackBerry. For Blackberry, the most influential factor is the price, the distribution channel is second high and the social factor is the lowest. For iPhone, the most influential factor for iPhone selection is a brand itself. The product factor is second high and the social factor is also the lowest.

In [5], Saranya Ruensri studied the behavior choice to buy mobile phone between blackberry and other mobile phones of students in Chiang Mai University. The information was of primary source coming from questionnaire interview of 400 samples from total 24,592 students during May 2010. It was found that product price, buying objectives, functional capability of the product, and the current popularity of specific mobile phone models were associated with the decision to buy Blackberry or mobile phone. Functional capacity of the product factor is the most influential factor.

3. Research Methodology

There are three main objectives of this paper:

1) Puspore of tablet usage is obseverd.

2) Various factors affecting the selection of tablet are presented. The priority of various factors affecting the use of tablets is also investigated.

3) Investigate the need in tablet usage between iPad and Samsung Galaxy Tab for two different population clusters.

3.1 Population

There are two different population clusters. First population cluster is students of Rangsit University.There were 400 sample students from total 25,424 students. The other population cluster is the Army Aviation Center's Pilot. There were 175 samples from total 320 pilots.

3.2 Conceptual Framework

Independent variables:

Sample population:

- Students of Rangsit University
- Pilots of the Army Aviation Center's Pilot

Tablets:

- The new Ipad
- Samsung Galaxy Tab7.7

Dependent variables:

- Purpose of tablets usage
- Factors affecting to tablets selection
- tablets selection

3.3 Tool design

Before designing of the questionnaire, this paper has surveyed the target of tablet usage and other factors from sample students and pilots. From roughly survey of two population clusters, it was found that main target of tablets usage between students and pilots are different. Therefore, the questionnaires used to survey for two population clustersshould be different in some item. This paper designs two questionnaires used to be tools for survey. Both two questionnaires consists of 4 parts. Pilot questionnaires contains as follows:

- Personal information
- Readiness and Need in laptops or tablets
- Factor influencing the selection on tablet between iPad and Samsung galaxy tab
- Suggestion and comment

Student questionnaires contains as follows:

- Personal information
- Behavior on tablet usage
- Factor influencing the selection on tablet between iPad and Samsung galaxy tab
- Suggestion and comment

The main difference of two questionnaires in part 3 as follow:

Questionnaires for students: the factors affecting the selection of tablets as follow: product (functional capacity), price, place for buying, brand, seller, and the current popularity of specific mobile phone models.

Questionnaires for pilots: the factors affecting the selection of tablets as follow: product (functional capacity), price, and brand.

3.4 Collect Data

For Rangsit University's students, the questionnaires were done during September 2012 to January 2013. For pilots, the questionnaires were done duringDecember2012 to March 2013

3.5 Analysis Data

This paper uses statistic mathematic that are frequency, percentage, mean, standard deviation.

4 Experimental Result

There are two subsections: Rangsit University students and Pilots of the Army Aviation Center's Pilot.

4.1 Students of Rangsit University

1) Tablets Selection of Students

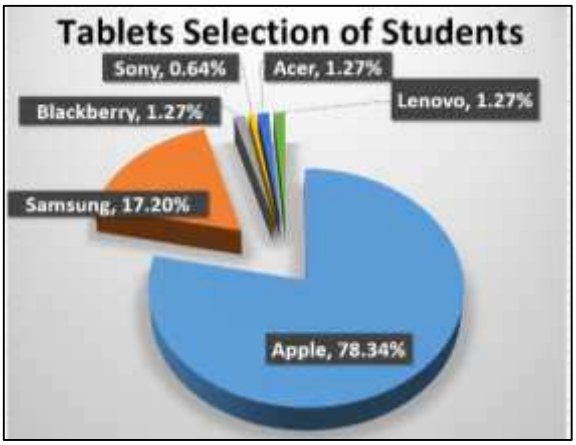


Figure 1. tablets Selection of Students

Figure 1. shows tablets selection of students of Rangsit University. The results found that in average, 78.34 % of Rangsit University's students preferred iPad, which it is the highest. Samsung Galaxy Tab is ranked as the second with 17.20 %. Finally, Sony tablets comes as the lowest in the list with only 0.64%.

2) Purpose of tablet Usage for Students

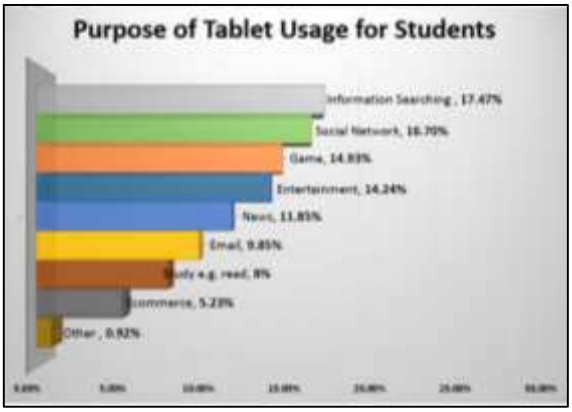


Figure 2. Purpose of tablets usage of Students

Figure 2 shows the target of tablets usage for students of Rangsit University. It is seen that the purpose of using tablet is mainly for information searching which can be accounted for 17.47%. Social Network is ranked as the second with 16.70%. Game is ranked as the third with 14.93%. Entertainment is ranked as the fourth with 14.24%

3) Factors affectting the selection of tablets for Students

Table 4.1 Factors affectting the selection of tablets between iPad and Samsung for students

Influential factors for selection	***iPad***		***Samsung Galaxy Tab***	
	mean	S.D.	mean	S.D.
1.Product *	4.37	0.49	4.00	0.60
2.Price	3.97	0.77	4.02	0.75
3.Distribution Channel	4.13	0.68	4.15	0.65
4.Brand**	4.45	0.63	4.03	0.70
5.Seller	4.20	0.70	4.03	0.74
6.Social Factor**	3.96	0.71	3.76	0.75

*It should be noted that product factor is functional capacity of product.**Brand factor includes a product style and option devices. ***Social factor includes the current popularity.

For iPad in table 4.1, it is seen that the most influential factor for iPad selection is a brand itself and the MEAN equals to 4.45. Product factor is ranked as the second with MEAN 4.37. Finally, Price factor comes as the lowest in the list with MEAN 3.97.

For BlackBerry in table 4.2, it is shown that the most influential factor for BlackBerry selection is the distribition channel and the MEAN equals to 4.15. the brand and seller factorsare ranked as the second with MEAN 4.03. Finally, Social factor comes as the lowest in the list with MEAN 3.76.

4.2 Students of Rangsit University

1) Tablets Selection of Pilots

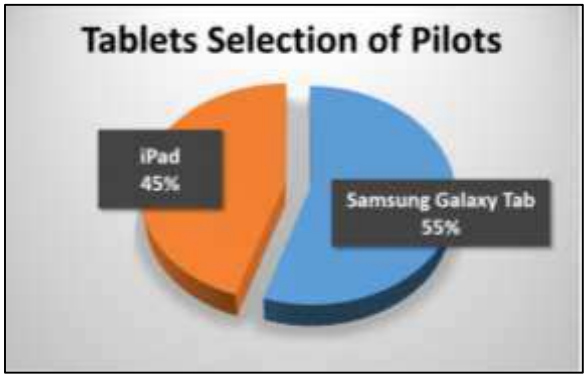


Figure 3. tablets Selection of Pilots

Figure 3 display tablets selection of Pilots The results found that in average, 55.00 % of Pilotspreferred Samsung Galaxy Tab.

2) Purpose of tablet Usage for Pilots

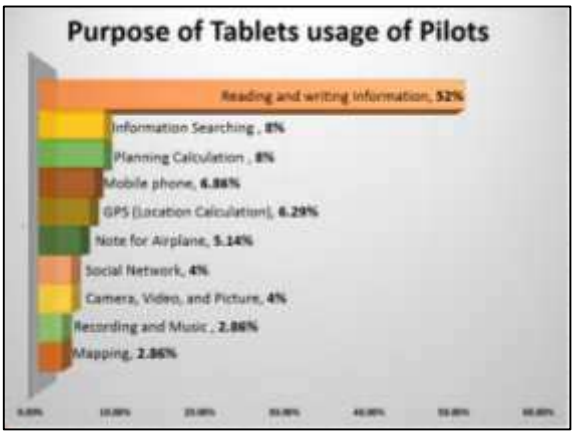


Figure 4. Purpose of tablets usage of Pilots

Figure 4 diplays the objective of tablets usage for pilots. It is shown that the purpose of using tablet is mainly for the Reading and Writing Informationwhich can be accounted for 52 %. The Information Searching and Planning Calculation are ranked as the second with 8 %. The Recording and Music and Mapping are ranked as the lowest with 2.86 %.

3) Factors affectting the selection of tablets for Pilots

Table 4.2 Factors affectting the selection of tablets between iPad and Samsung for pilots

Influential factors for selection	***iPad***		***Samsung Galaxy Tab***	
	mean	S.D.	mean	S.D.
1.Brand	3.72	0.74	3.85	0.74
2.Product*	3.70	0.73	3.87	0.73
3.Price	3.66	0.79	3.79	0.80

*It should be noted that product factor is functional capacity of product. **Brand factor includes a product style and option devices.

For iPad in table 4.2, it is seen that the most influential factor for iPad selection is thebrand and the MEAN equals to 3.72. the product factor is ranked as the second with MEAN 3.70. Finally, Price factor comes as the lowest in the list with MEAN 3.66.

For BlackBerry in table 4.2, it is listed that the most influential factor for Samsung selection is the product and the MEAN equals to 3.87. the brand is ranked as the second with MEAN 3.85. Finally, price factor comes as the lowest in the list with MEAN 3.79.

5.Conclusion and Discussion

5.1 Conclusion

For students, the results found that in average, 78.34 % of Rangsit University students preferred iPad to Samsung Galaxy. Interestingly, the purpose of using tablet is mainly for information searching which can be accounted for 56.75%. The most influential factor for iPad selection is the brand itself and the MEAN equals to 4.45.

For pilots, pilots give a higher priority of iPad than Samsung in all the factors affecting selection tablets as listd in table 4.2. From the result, it is found that the functional capacity of the product factor is the most influential factor. Interestingly, the target of using tablet is mainly for reading and writing information which can be accounted for 52.00%. In average, 55.00% pilots preferred Samsung Galaxy Tab to iPad. It is interesting to note that the different population clusters are different the purpose of using tablet and the requirement for tablet selection

5.2 Discussion

Students of Rangsit Univesrity is prefferd iPad, which is the same direction as the research of Piper Jaffiay Inc. on April 2013 that iPad is the most popular for american teenager [6].

However, the main purpose of tablet usage of the thai students in this paper is the information searching while the main purpose of american is for playing game researched by Admob in 2011[7].It is should be noted that the target of the information searching of american is ranked as second.

Pilots prefer Samsung Galaxy Tab to iPad since the main purpuse of tablet usage is for working that is to read and write infromtaion during flying. Samsung Galaxy Tab is easily enable transfering data from other devices to tablets while iPad is limited which is the same as

in[8].

For studen maket, Ipad is preferred and purpose of using tablet is mainly for information Searching. This means that samsung should improve his product in this feature.

For worker maket, Samsung is preferred and the purpose of using tablet is mainly for reading information. This means that Ipad should improve his product in tranfering data form other devices to tablets.

5.3 Future work

The population clustershould be extended to other clusters in order to get more information.

Acknowledgment

Authors many thank to LT.COL.Siripong Phurat for collecting information from pilots of the Army Aviation Center's Pilot. Authors also a big thank to Ubonrat Suksirikulfor collecting information from students of Rangsit University.

References

[1] Kelly Hodgkins, United and Continental Airlines to hand out 11,000 iPads to its pilots. http://www.tuaw.com/2011/08/23/united-and-continental-airlines-to-hand-out-11-000-ipads-to-its/, Aug 23rd 2011.

[2] Menapa and Sanon Chimmanee, Teachers and parents' readiness for supporting tablet use in 1st grade classroom of Chanthaburi Primary Educational Service Area Office 1, Thailand, BUU international conference 2013 at Jontiem Plam Beach, Pattaya, 4-5 July

[3] John Laugesen, and Yufei Yuan, What Factors Contributed to the Success of Apple's iPhone, IEEE Computer Society, Proceeding of ICMB-GMR '10 Proceedings of the 2010 Ninth International Conference on Mobile Business / 2010 Ninth Global Mobility Roundtable, Pages 91-99, 13-15 June 2010, Athens

[4] Jutharut Kunasakpanich, A study of behavior and factors affecting the selection of the smart phone between iPhone and Backberry of Rangsit University's students, Independant Study (IS), Asst.Dr.Sanon Chimmanee is adveror, M.S.ITM-Online, Rangsit Unversity in 2012

[5] Saranya Ruensri, The comparison of behavior choice to buy moble phone between BlackBerry and other moblie phones of students in Chiang Mai University, Independant Study (IS), Master of Economics, Chiang Mai University in 2010.

[6] Piper Jaffray, Taking Stock With Teens, Spring2013http://www.whiteboardadvisors.com/files/Taking_Stock_Teach-in_Spring_2013_MV_2.pdf

[7] AdMob/Google, Understanding Tablet Device Users, March 2011, http://www.thinkwithgoogle.com/insights/uploads/17683.pdf

[8] Apirak hanpichitvanit and Pichet suwiyanon, The New iPad, 60vi, 29 May 2012, http://www.60vi.com/clip/the-new-ipad/

ACIS 2013

Asian Conference on Information Systems 2013

Special Session on Service Science

A preliminary proposed curriculum support system for information technology specialists based on the service science approach

DONG, Wooseok†* **SHIRAHADA, Kunio†** **KOSAKA Michitaka†**

†School of Knowledge Science School
Japan Advanced Institute of Science and Technology
s1160005@jaist.ac.jp

Abstract

Legacy curriculum support systems have simple functions for searching syllabi and course registration. However, students, who are educational service customers, have different characteristics. Nevertheless, students are in most cases provided fixed curricula designed solely by educational service providers.

In this paper, we propose a preliminary curriculum support system for information technology specialists based on a service science approach. The aim of this system is to support completion activities through which students can design a suitable curriculum to reach to their individual career goals. In the design step of the system, we gathered data from 30 students using the persona strategy method. We then, undertook a case study by redefining the education service of University A based on IEEE&ACM's Computer Curricula 2005 report. We also applied knowledge space theory to the system interface for visualization of ideal, current and increased knowledge.

For evaluation, we selected 10 graduates of University A. We checked the correlations between the curriculum provided to these graduates and the curriculum recommended by the system. As a result, the system recorded a 57.66% correlation between the recommended courses and courses taken by graduates. We expect that this system can become an interactive support system between students and service providers for value co-creation with the role of designing suitable curriculum by students.

Keywords: Service science, Information technology, Curriculum support system, Knowledge space theory, Computing Curricula 2005

1 General Instructions

Recently, the information industry has become increasingly complicated with the rapid advance of information technologies; therefore, educational service providers of information science have to consider the question how to educate information technology specialists [1]. Multiple potentials such as communication ability; problem solving and project management skill are required for working. To meet these needs, the Computing Curricula was published by IEEE&ACM for guidance of graduate schools; many universities design education program based on these guidelines.

However, legacy curriculum support systems provide only simple functions for searching syllabi and course registration. These systems do not consider the characteristics (career goals, interests and potentials) of registrants [2]-[4]. Educational service providers design fixed curricula based on their roles on the education program, and provide them to students like consumer products. Thus, students cannot receive suitable curricula for accomplishing their career goal from legacy curriculum support systems.

In terms of the service science approach [5]-[7], legacy support systems are Goods Dominant Logic based. On the other hand, if curriculum support systems are designed in accordance with Service Dominant Logic [8], educational service providers can offer suitable processes of completion activities for value co-creation by offering a suitable curriculum designed by students in cooperation with educational institutions.

To test this idea, we designed a curriculum support system based on Service Dominant Logic concepts and persona strategy method [9]. This system aims at supporting completion activities based on the relationship between career goals and courses. We considered standards for display

of interactive activities processes and co-creation value evaluation. We then applied knowledge space theory to the interface of our system to display completion activity processes and the courses required to reach career goals for knowledge extension based on students` current state of knowledge.
In addition, we had to devise mutual standards for the content and scope of knowledge and competencies required to reach career goals. For this purpose, we adapted the findings of the Computer Curricula 2005 report to the system.

In this paper, we introduce the prototype implementation case study conducted at the information science school of university A.

1.1 Computing Curricula 2005

IEEE&ACM Computer Society jointly sponsor the development of a Computing Curricula volume on Computer Science [10]. These volumes have helped to set international curricular guidelines for graduate programs in computing. Computing Curricula 2005(CC2005) presents five kinds of degree programs; CE (Computer Engineering), CS (Computer Science), SE (Software Engineering), IT (Information Technology), and IS (Information Service).

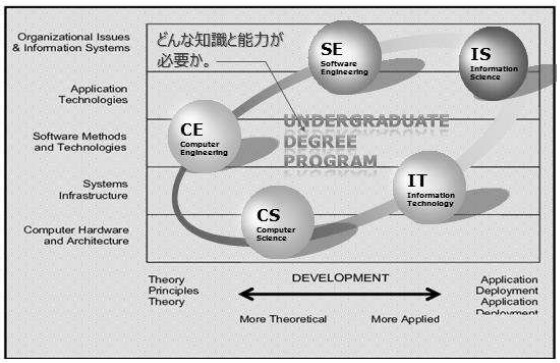


Figure 1. Five degree programs of CC2005

The CC2005 is guideline for curriculum with comparative weight (from 0 to 5 points) and relative performance capabilities (from 0 to 5 points) over 11 areas in not only computing courses but also non-computing courses for graduate degree programs. We applied the CC2005 report to proposed system for three reasons: it was published by the International Committee of IEEE and ACM Computer Society; it can be widely applied to other information science universities and colleges; it allows a relationship between information science industries and information science education programs to be mapped.

1.2 Service science approach

When educational service providers design education programs, CC2005 will be a guideline. According to CC2005, we can present standards to students of what knowledge required reaching career goals with the five degree programs. In addition, we can define knowledge level based on comparative weight and relative performance capabilities. Additionally, CC2005 provides a guideline about suitable curriculum to students with specific career goal. Given that information technology education is a service, the value of this service is important to both service providers and students.

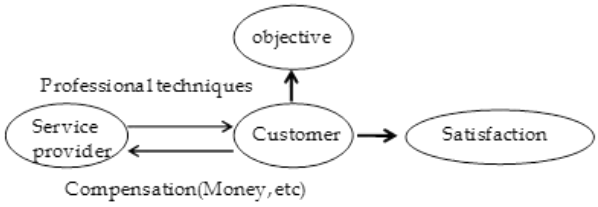


Figure 2. Definition of Service

According to Figure 2, service providers have to provide service to customers based on professional techniques and gain compensation. Moreover, service providers have to consider customers, objects and satisfaction. That means service providers have to consider increasing service value and customers' satisfaction [11]. In addition, service providers have to provide service based on the customers' (students) objects (purposes) and characteristics. In the proposed system, we define service value as follows: students can take suitable courses and curriculum with individual characteristics for achieving their career goals. Under the proposed system, students can require the service providers to provide something these needs and attempt to meet them. As a result, students can acquire service value and feel satisfaction with the service provided from the educational institution [12]-[15].

1.3 Concept of Knowledge Space Theory

Actually, education service value can be defined differently according to situations, institutions

and targets. In this paper, the definition of value is that students accomplish career goals.
For co-creation of value, basis for a mutual understanding of applicable standards is required among service stakeholders. We defined career goals with required knowledge and knowledge level for acquire process with expansion of knowledge [16].
Hence, we re-composed the curriculum of University A based on the integration CC 2005 with knowledge space theory as a user interface for supporting completion activities.

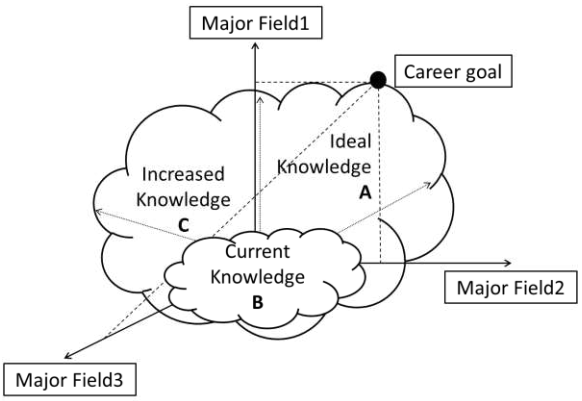


Figure 3. Concept of Knowledge Space Theory

Knowledge space theory was proposed in 1985 by Jean-Paul Doignon and Jean-Claude Falmagne to explain the formation of a given province of knowledge. The formal concept is that a field of knowledge is specified by a finite set of items, i.e., problems or tasks a student may or may not be able to solve [17]. The set of all possible knowledge states is called a knowledge space.
In this paper, we identify career goals a student may be able to achieve. Each student can attempt to reach them from his or her current knowledge state. Therefore, we can show a goal-related degree program and courses to students in a knowledge space when students select career goals (ideal knowledge A) and current knowledge level (current knowledge B) as axes. In additional, we can identify subjects that are required knowledge for reaching career goals (increased knowledge C) as expanding from current knowledge B to reach ideal knowledge A. Knowledge space B expands when students take a course, do research, or get credentials or degrees. We expect to stimulate students' motivation for completion activities by visualization of goals, knowledge and potential. In addition, we anticipate gathering data about student's activities and changed knowledge space with career goals for service innovation of information technology education.
Service providers can support students' career goals and provide new standard guidelines for completion activities based on analysis of gathered data. Students can design curricula and undertake completion activities. When students' knowledge space expands, students can visualize information and impressions with service stakeholders based on the proposed system.
As a result, we expect service value co-creation based on interaction between students and service providers using the proposed system.

2 Methods

In this chapter, we explain the curriculum design process and the method of creating a knowledge space with a case study of University A's school of information science [18].

2.1 Curriculum re-composition

To create a curriculum structure in the knowledge space, we divided all of University A's courses based on five degree programs and areas of the CC2005 guidelines. In addition, we mapped the relations among courses [19].

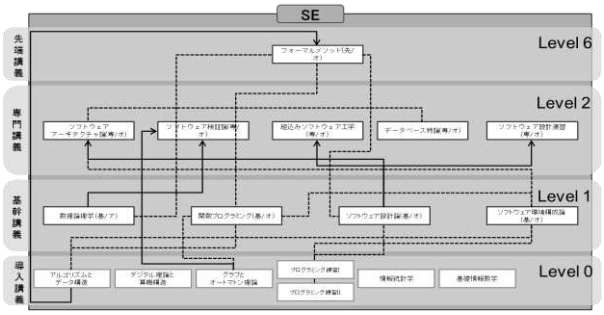


Figure 4. Relationship diagram of University A (software engineering area)

As a result, we drew a relationship diagram of courses based on CC2005, as shown in Figure 4. All subjects were reclassified based on CC2005. We extracted the subject areas need to reach career goals by applying the knowledge space of each axis as defined in Figure 6.

2.2 Design of knowledge space

2.2.1 Definition of career goals and knowledge axe

First, we have to develop a career goal list and a knowledge axis with subjects (major field) for knowledge space creation.

Major Field	Subject	Major Field	Subject
M_1	Algorithm	M_8	Database
M_2	Operating system	M_9	Software Foundation
M_3	System Development	M_{10}	Network & Communication
M_4	Artificial Intelligence	M_{11}	Computer Graphics
M_5	Logic and Mathematics	M_{12}	IT management
M_6	Human Interface	M_{13}	Embedded System
M_7	Hardware and Device	M_{14}	Security

Figure 5. List of subject fields based on CC2005 and A University program.

We identified the 16 career goals (G1, G2 …Gn) below based on a survey of 30 students from University A using the persona strategy method.

In addition, we added 2 specialist subject areas (embedded systems and security) that are provided by University A to the areas on the CC2005 list. 14 subject areas (Major Field, M1, M2 … Mn) were defined as axes of the knowledge space. (Figure 5)

We made the table shown in fig.6 considering University A's information science school to define the knowledge space.

Then, to each career goal, 4 major fields were substituted as knowledge space axes (A1-A4) based on high comparative weight and relative performance capabilities as defined in the CC 2005 guidelines. (Figure 6.)

2.2.2 Definition of knowledge level

To define the knowledge level, we specified relative performance capabilities in subject field from a minimum of 0 to a maximum of 8.

We suggested from a minimum of 0 to a maximum of 5 for M.S. candidates and from a minimum 5 to a maximum of 8 in a subject area for Ph.D. candidates. The A1 axis below shows the courses with the highest comparative weight in the CC 2005 guidelines.

No.	Categories of Goal Career	CC Field	Knowledge Axis			
			A_1*	A_2	A_3	A_4
G_1	Development of hardware and Device	CE	M7	M1	M2	M3
G_2	Design of human interface	CE	M6	M1	M3	M7
G_3	Research about Operating system	CE	M2	M1	M9	M7
G_4	Development of embedded system	CE	M13	M7	M3	M2
G_5	Design an algorithm	CS	M1	M4	M5	M11
G_6	Development of artificial intelligence	CS	M4	M1	M6	M5
G_7	Research about Computer graphics	CS	M11	M1	M6	M5
G_8	Design of logical model and mathematical theory	CS	M5	M1	M2	M9
G_9	Software development	SE	M3	M8	M10	M2
G_{10}	Research of software foundation	SE	M9	M3	M2	M5
G_{11}	Management of development	SE	M12	M3	M8	M10
G_{12}	Management of embedded system development	SE	M13	M3	M12	M7
G_{13}	Research about network	IT	M10	M1	M5	M2
G_{14}	Management of network	IT	M10	M12	M7	M2
G_{15}	Research about security	IT	M14	M1	M7	M10
G_{16}	Management of Information business or service	IS	M12	M3	M8	M10

Figure 6. Setting of knowledge axis

On the other axes (A2-A4), we will provide optional functions for students to adjust knowledge level by themselves.
In the first step, current knowledge B is created by initial value (from 0 to 2) based on input data from students background, information science experience and licenses or credentials.

2.2.3 Knowledge space creation

Figure 7 shows a flow-chart that explains the knowledge space creation.
Steps 1 and 2 are defined in figure 6.
In step 3, students can choose a career goal from the categories of career goals on the proposed system. Next, in step 4, the system identifies 4 subject areas related to career goals based on figure 6. Then, in step 5, the system recommends the ideal knowledge space (as explained above) and creates the current knowledge space.
In step 6, students can choose a curriculum from subject areas (shown in step 4) on the system.
Students undertake completion activities based on a curriculum from step 6. As a result, professors give credits and comments to students.
Finally, completion activities results and information are input into our system by professors, and the current knowledge space is updated automatically.

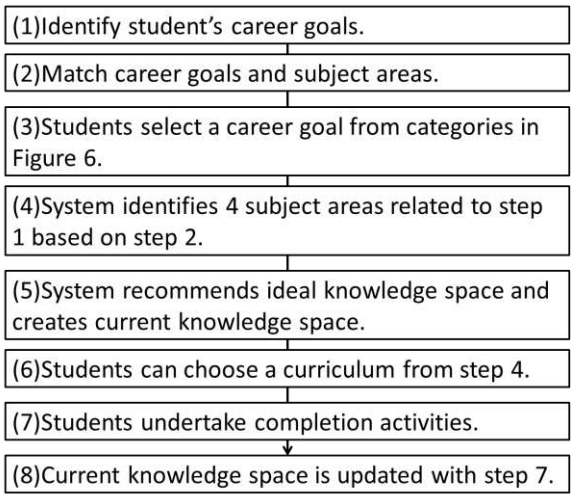


Figure 7. Knowledge space creation flow chart

For example, to create a knowledge space for M.S students who want to engage in algorithm research, the following procedure applies: Students choose their career goals from G5 (including CS degree program shown in figure 6); the A1 axis is then set to the Algorithm (M1) subject area and ideal knowledge level 5 automatically.
In addition, Artificial Intelligence (M4), Logic and Mathematics (M5) and Computer Graphics (M11) are selected as relative subject areas. If the student selects Artificial Intelligence (M4) as a sub research area in subject areas, M4 becomes the A2 axis.
Finally, the knowledge space is created as follows;

KS(MasterStudent , G5)={A1*=M1(5), A2**=M4(4),A3=M5(3),A4=M11(3)}

2.3 Method of increasing knowledge

The relationship diagram shows how students reach goal levels on the knowledge space (figure 8). On the A1 axis, the current knowledge level is 1 and the ideal knowledge level is 5, so increased knowledge variance between ideal and current knowledge level can be calculated. Figure 8 shows the completion process in the algorithm area based on the knowledge space. This student is recommended to take courses on level 2 and 3 layers.

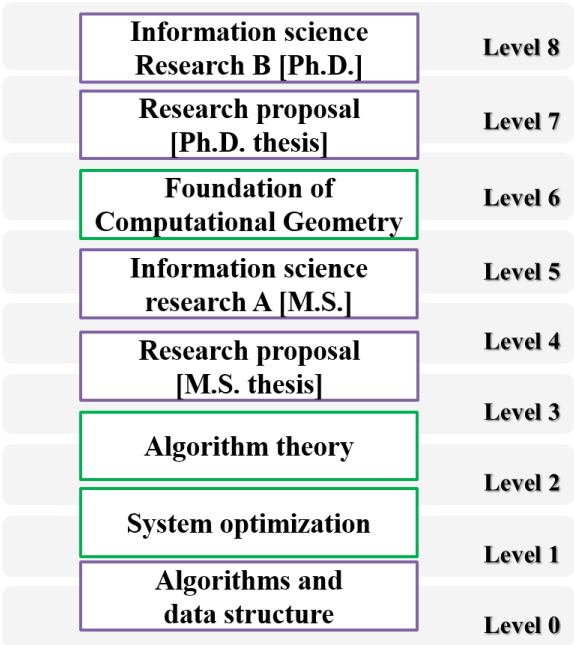


Figure 8. Interface of the knowledge space

The student then has to make a research proposal on the level 4 layer, and to write an M.S. thesis on the level 5 layer. Other axes can be represented through this method.

3 Implementation

In this chapter, the usage scenario is explained from the user's viewpoint.

3.1 Usage Scenario

To utilize this system,

1) students input personal data (profile, education background, language ability) as basic data.
2) students create a new current knowledge space based on the input data.
3) students select a course (M.S. / Ph.D.) and career goal from 14 goal careers listed in the box.
4) the proposed system creates an ideal knowledge space with 4 axes.
5) the proposed system selects a main-research subject area axis automatically, and students select a sub-research subject area from the 3 remaining subject areas axes.
6) students make a curriculum on the system interface (figure 8) together with an academic adviser. After making the curriculum, required conditions for graduation are checked automatically.
7) students undertake completion activities following the curriculum on step 6; academic advisers input educational records with comments into the system.
8) during completion activities, the current knowledge space is updated with educational evaluation from academic advisers.

In this paper, we implemented a prototype system based on scenario from step 1 to step 6. Step 7 forward will be considered in future research.

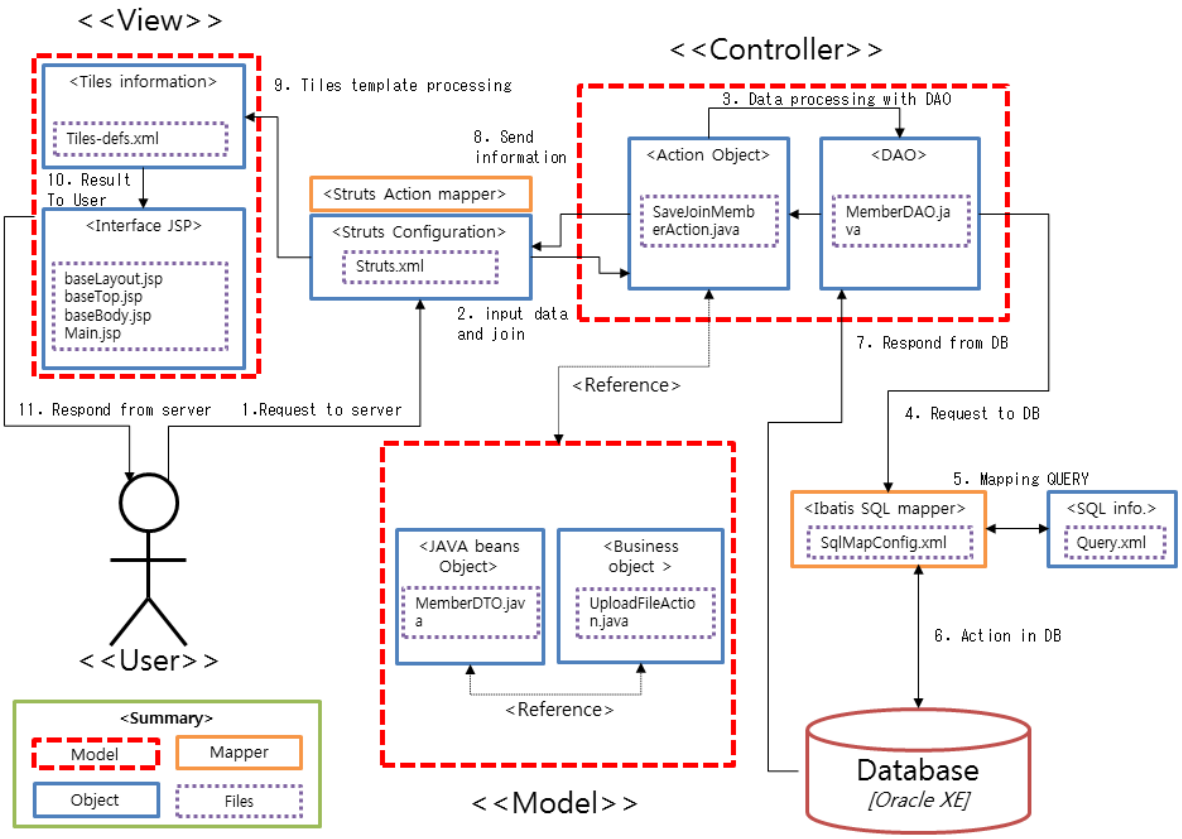


Figure 9. Integration diagram of usage scenario

3.2 User Interface

We developed a prototype system with layer interface (figure 8) of the knowledge space.
In figure 9, when student selects the name of the axis for a major field, the current knowledge layer (blue color) and ideal knowledge layer (red color) are displayed. On the layer structure, courses which are vertically related are “core” courses; these which are horizontally related are “elective” courses. Each lecture is selected by clicking.
Moreover, course information (unit, lecturer, difficulty) can be shown on the interface.
Additionally, required conditions for graduation can be checked automatically.

4 Evaluation

To evaluate this system, we select 10 graduates from university A who gave us feedback on this system. We are gathered personal information and data from the interviewees. In addition, we compered the graduates’ curricula and the system’s recommended curriculum for correlation.
The results are shown in figure 10. The average correlation is 57.66%. The highest point recorded was 76.8%, in the Information Service area. The Computer Science area and the Software Engineering area also recorded over 60% correlation. Correlation in the Computer Engineering area was around 50% correlation and the Information Technology area recorded the lowest correlation, around 40%.

ID	Goal Career	Goal Ability	Accorded Lecture		Probability
			Accord	Total	
s1000001	M Program	Development of hardware and device	5	9	55.5%
s1000002	M Program	Research about Operating system topic	6	12	50%
s1000003	M Program	Development of Software	4	7	57.1%
s1000004	MD Program	Management of embedded system development	7	11	63.6%
s1000005	D Program	Management of information and business process	10	13	76.8%
s1000006	M Program	Management of Network issues	5	10	50%
s1000007	M Program	Research about network security issues	6	11	54.5%
s1000008	M Program	Research about Network issues	4	10	40%
s1000009	M Program	Design of human interface	5	8	62.5%
s1000010	D Program	Design of artificial intelligence	6	9	66.6%
Total Probability					57.66%

Figure 10. Evaluation result

The average correlations were lower than our initial predictions. Moreover, we discovered that “core” courses’ correlation was over 90% but “elective” courses’ correlation was only around 20%.
From interview with graduates, we learned that is students most graduates had selected courses to get credit easily or based on other students’ advice. However, the interviewees felt that the proposed system would have been helpful in choosing courses when they were students.

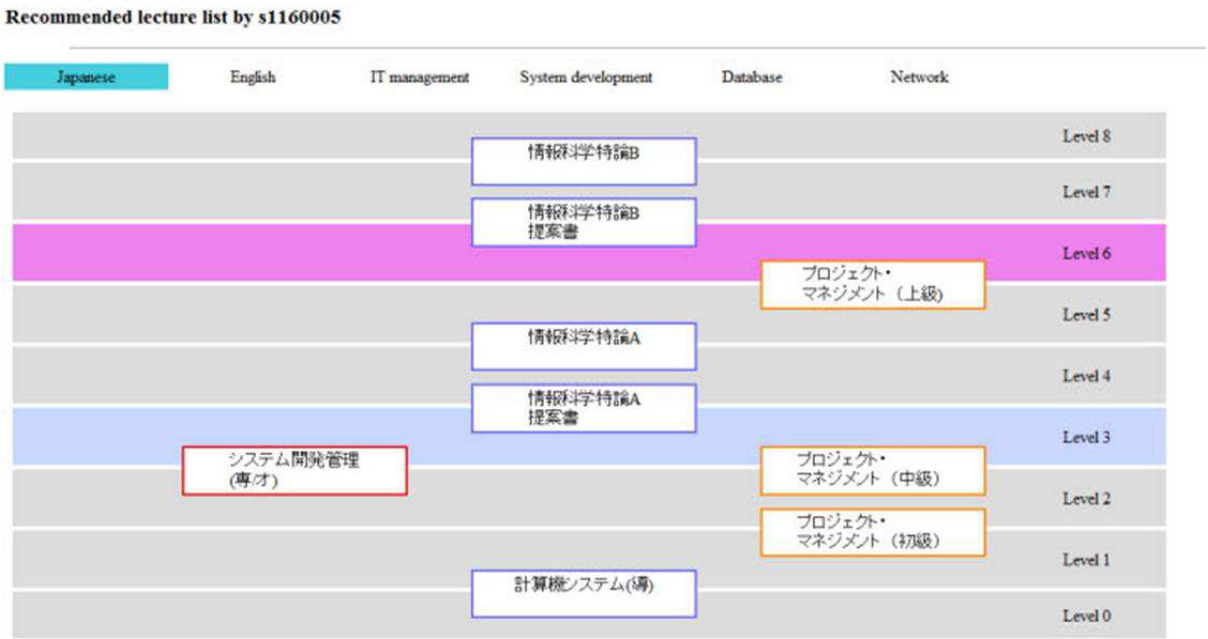


Figure 9. Curriculum design page interface

5 Discussion

According to interviews and feedback, students had positive reactions and comments about the knowledge space interface and curriculum visualization with student characteristics. In addition, additional functions that allowed them to check required conditions for graduating were also evaluated positively.

However, concerns exist about the major field, courses and knowledge level. Defining career goals and major fields at this point encounters a problem: how to generalize the career goals and the major field this system to the curricula at other institutions. In particular, freedom of selecting subject areas and axis was evaluated poorly.

Despite that, proposed system takes a service approach that considers not only service from provider to customer, but also customer to provider and creates service value through cooperation and co-creation. We have a plan for future work that how to provide a balanced curriculum which 'core' and 'elective' courses to student.

We hope to that, after completion of development, this system becomes the interactive supporting system platform between students and service providers for value co-creation in not only information technology education but also other areas of education.

Acknowledgment

The authors would like to thank Professor K.Ochimizu, Ph.D. A. SEO and Professor Wm.R.Holden III for their valuable advices.

References

[1] O. Koichiro, New IT specialist education in Ishikawa prefecture，ISICO (2010)

[2] Fujimoto T, Hiraishi K, Miura M, Kunifuji S, Proposing and designing courseware syllabus system to assigned students in research laboratory, IPSJ, (2006)

[3] Takahashi K, Hori Y, Imai Y, A Course Registration Support System based on Syllabus Analysis and Spreading Activation Model, The Institute of Electronics, Information and Communication Engineers, (2008)

[4] Yarita T, Miyoshi T, Curriculum-Based Study Support System, The Institute of Electronics, Information and Communication Engineers, (2006)

[5] Akio, Kameoka, Service science, *NTS*, 2007

[6] C. Lovelock, L. Wright, The principle of service science, *Hakutou press*, 2001

[7] B. Stauss, K. Engelmann, A. Kremer, A. Luhn : Services Science, *Springer* (2008)

[8] R. F .Lusch and S. L. Vargo ：The Service Dominant Logic of Marketing, *ME. Sharpe, Inc*. (2006)

[9] J. Pruitt and T. Adlin ：Persona marketing, *Diamond press*, 2007

[10] IEEE&ACM，Computing Curricula 2005, The Overview Report, ACM&IEEE, (2005)

[11] K Zhang, M Kosaka, K Shirahada, T Yabutani, A Proposal of B to B Collaboration Process, *IEEJ Transactions on Electronics*, Information and Systems 132(6), 1035-1040, 2012-06-01

[12] B. Joseph PINE II and James H. Gilmore：The Experience Economy，*Harvard Business School Press* (1999)

[13] R.J. Glushko："Designing a service science discipline with discipline", I*BM System Journal*, http://www.research.ibm.com/journal/sj/471/glushko.htm

[14] Fisk, Grove, John：Interactive Services Marketing，*Houghton Mifflin* (2008)

[15] Jon Sundbo and Per Darmer：Creating Experiences in the Experience economy，*Edward Elgar Publishing* (2008)

[16] M Kosaka：The approach to the model of knowledge growth，*Syakaishouron press*, 2010

[17] H.M.Belal, Kunio Shirahada and Michitaka Kosaka : Knowledge space concept and its application for servitizing manufacturing industry, Journal of service Science and Management, 2012

[18] DONG Wooseok, SEO Akaki, Ochimizu Koichiro, The research of completion supporting system for education program through persona strategy,IPSG(2011)

[19] JAIST DATA BOOK and syllabus 2010, JAIST, www.jaist.ac.jp, (2010)

How do Sharing Service Providers Create Value?

Youji Kohda† **Kazuhiro Masuda†**

†School of Knowledge Science
Japan Advanced Institute of Science and Technology, Nomi, Japan
{kohda, keith-masuda}@jaist.ac.jp

Abstract

The behavior of customers has changed from total consumption to collaborative consumption in the past decade. Sharing services are now becoming a major part of our daily lives. From the view of customers, they have both benefits and risks. They offer benefits such as saving money and protecting the environment, but they also have non-negligible risks, such as the risk of becoming involved in injury cases in the worst-case situation. Therefore, it would make sense for people to choose not to use sharing services since human beings are risk averse by nature. Nevertheless, why do people choose to use sharing services? In this paper, we present a hypothesis to answer this question. We conjecture that a successful service provider always bears the sharing risks that otherwise would be borne by the customers. Thus, the customers of sharing services could reap the benefits of sharing, and this explains the recent rise in the popularity of sharing services. We also restate our conjecture from the view of service dominant logic, especially in the context of value co-creation.

Keywords: service value, sharing service, risk and benefit, risk management, value co-creation

1 The service age

The Internet has changed and is still changing our lifestyles dramatically. For example, let us consider making a hotel room reservation. Just a short time ago, you had to make several phone calls until you found a vacant room because there was no way to see whether a room was available before you called the hotel. Now the Internet has changed the situation completely. You can see the availability of hotel rooms on an online map, and all you need to do is pick one from the map (Figure 1). This is a simple but illustrative example of service in the Internet age. The development of ICT[1] is a key driver for the skyrocketing growth of services in the Internet. Our current world cannot exist without ICT.

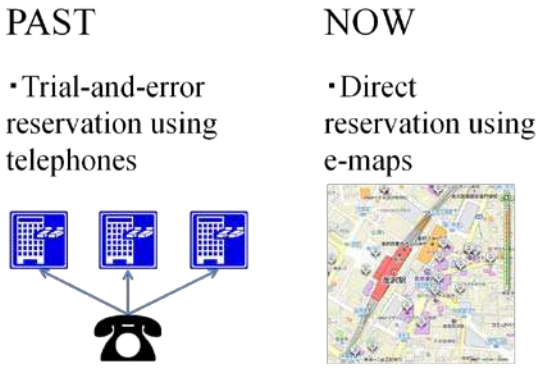


Figure 1. Change of way of making hotel room reservation

As shown in Figure 2, there were just a few technology stacks in 1995. You had to prepare the rest of the technologies by yourself to start even just a simple service. Now that ICT has developed enough, you can start your own service at low cost or even free of charge. Cloud computing is accelerating this change. You could even start your own business today. This means that we are in the age of idea first, technology second.

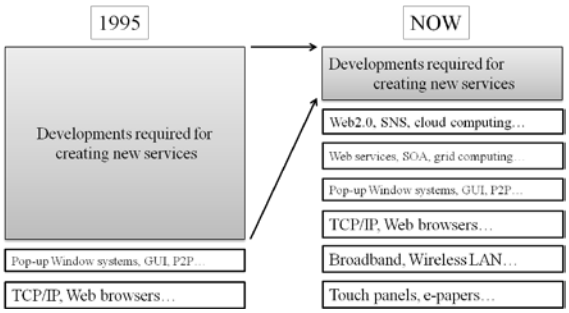


Figure 2. Changes in service infrastructure

[1] ICT stands for Information and Communications Technology

The development of the Internet has opened the door to a new movement. Sharing services are changing our lifestyles slowly but surely [1]. This movement will affect not only our personal lives but also business enterprise activities because it will force enterprises to transform their business models into consumption economics-based business models [2].

This paper discusses sharing services in detail. We seek the reason for the increasing popularity of sharing services. We also restate this from the view of service dominant logic, especially in the context of value co-creation [3].

2 A new trend: sharing services

2.1 Collaborative consumption

The behavior of customers has changed from total consumption to collaborative consumption in the past decade. Sharing services are now becoming a major part of our daily lives.

In the 20th century, people were competing for the ownership of brand new products, but at the beginning of the 21st century, people have just been learning that this way of living is no longer sustainable. This changing of the minds of people is advancing the penetration of sharing services in our daily lives.

2.2 Sharing services

Table 1 shows some real-life examples of sharing services. The first column shows the sharing service, and the second column shows the shared object of that sharing service. The third column shows the sharing type, either simultaneous use or sequential use.

For instance, for room/house sharing, the shared object is immovable estate (room, flat, house,…), and the sharing type is simultaneous use when, for example, a group of people live together in the same room, flat or house. For another instance, for car sharing, the shared object is moveable asset (car), and the sharing type is sequential use. People can only borrow cars when they are not reserved and are not being used by others.

In ride sharing, a car owner rides with other people to make use of unoccupied seats. In money sharing, a person with some money lends the money that is not being used directly to the person who really needs the money, understanding that the borrower needs the money.

Table 1. Sharing service examples

Sharing Service	Shared Object	Sharing Type
Room sharing/ House sharing	Immovable estate (room/flat/house)	Simultaneous use
Car sharing	Movable asset (car)	Sequential use
Ride sharing	Means of transportation (car)	Simultaneous use
Money sharing	Means of purchase (money)	Sequential use
Knowledge sharing	Means of problem solving (knowledge)	Simultaneous use
Skill sharing	Means of problem solving (staff)	Sequential use

In knowledge sharing, a person transfers his or her expert knowledge to another person who has a problem to be solved. Solving problems with open source development is an example of knowledge sharing. In skill sharing, a person works with another person who has a task to be finished. So-called pro bono work is an instance of skill sharing where a worker undertakes voluntary work without pay, for example, on weekends or public holidays.

We chose two representative sharing service providers: Zipcar[2] and Airbnb[3].

Zipcar is a leading car sharing service in the US. It is a self-service car rental service for people to borrow a nearby car as soon as possible. It is easy to borrow a car if you are a member of Zipcar (Figure 3). You can reserve the nearest available car using your smartphone and obtain the map that indicates the location of the car. You can even unlock the car door using your smartphone.

Airbnb is the leading room sharing service in the world. It matches hosts and guests, and the accommodation ranges from apartment rooms to

1. Reserve car.
2. Go to location.
3. Find car.
4. Unlock it and get in.

Figure 3. Steps for borrowing car by using Zipcar

[2] The “zip” in “zip-car” means “speed.”
[3] The original name was “AirBed&Breakfast.”

mansions and even castles. Note that this type of sharing is not always strictly simultaneous because the hosts may not stay with the guests for the period of stay. The hosts may lend their rooms or houses to the guests (tourists) while they are gone.

It is easy to find accommodation if you are a member of Airbnb (Figure 4). You first need to post your profile including your face photo. After doing this, you can find and book adequate accommodation.

3 Anatomy of sharing service

3.1 Risks and benefits

From the view of customers, a sharing service has both risks and benefits. Table 2 lists up the risks and benefits of Zipcar and Airbnb. As you can see, sharing services have non-negligible risks, such as the risk of becoming involved with injury cases in the worst-case situation. Thus, it would make sense for people to choose not to use sharing services since human beings are risk averse by nature.

Then, why do people choose to use sharing services? We conducted an interview to find out the answer to this question.

3.2 Managing risks

We interviewed one of management personnel of Zipcar, and it was found that Zipcar had intentionally built up a customer support infrastructure to reduce the risks of sharing [4].

The infrastructure is a combination of guidelines[4] and call centers. It helps the company to keep good-mannered customers by filtering

1. Post your profile with face.
2. Find place.
3. Book place.
4. Check into place.

Figure 4. Steps for reserving a place to stay by using Airbnb

[4] Zipcar's guidelines consist of "six simple rules." See http://www.zipcar.com/is-it/rules.

Table 2. Risks and benefits of sharing

	Zipcar	Airbnb
Risks	• Cannot make reservation for desired time • Car is not there even though you have made reservation • Car is dirty	• Cannot make reservation for desired time • Becoming involved in injury cases
Benefits	• Low cost compared to cost of private car • Sense of satisfaction due to contributing less to global warming	• Low cost compared to cost of conventional commercial accommodation • Chances to meet people and to have good communication

out bad-mannered ones who make the cars dirty or return the cars late. The members of Zipcar are gently requested to report any problems to a call center[5]. Thus, the call centers can discriminate between good-mannered and bad-mannered members (Figure 5). The important point is that the reports are almost mandatory because the customer will be falsely charged for a dirty car and/or late return of a car if they do not report the problems.

3.3 Remaining benefits

The benefit for good-mannered customers is that they can avoid encountering a dirty car and/or late return of a car because bad-mannered customers will have their membership revoked and will no longer be able to borrow cars. This is one of the service values that Zipcar is implicitly

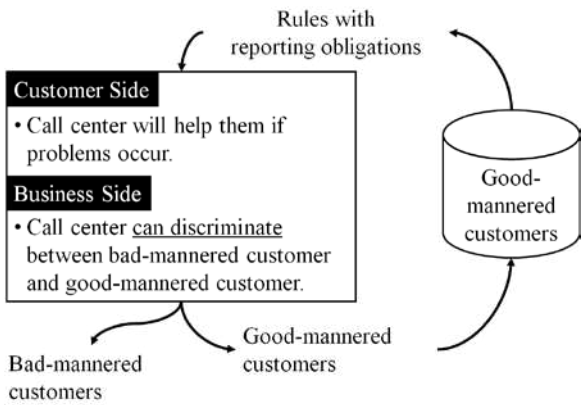


Figure 5. Rules with reporting obligations

[5] One of the six simple rules is as follows: "Before you drive away, inspect your Zipcar inside and out. To report damage, a dirty car or low fuel, call 1-866-…"

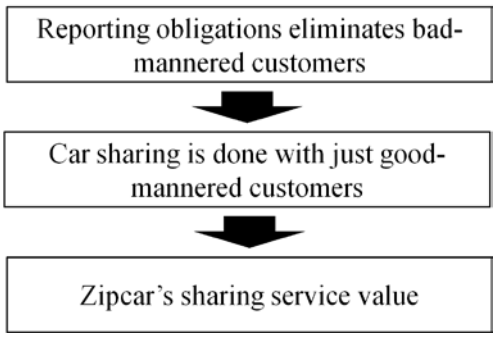


Figure 6. Logic behind ZIPCAR's success

offering to its customers (Figure 6).

From the service provider's view, this means that the company should invest a lot of money to bear the sharing risks that otherwise would be borne by the customers. Thanks to this investment, the customers of a sharing service can reap the benefits left of sharing solely.

4 Value co-creation in sharing services

4.1 Eliminating sharing risks from customers

A sharing service consists of several components, and thus the customer's experience of the service is affected by all the components. This means that the value of a service could be a (weighted) mean value of all the values of the components. This implies that you could increase the value of a service if you increase one of the components that has a smaller value than the mean. This was the idea behind the success of Zipcar: it has invested in infrastructure to reduce the number of dirty cars and cars returned late.

By generalizing this, we conjecture that *a successful sharing service provider always bears sharing risks that otherwise would be borne by the customers.*

4.2 Value co-creations in sharing services

Next, we think about the business models of the successful sharing service providers from the view of service dominant logic, especially in the context of value co-creation.

Value co-creation is "a resource integration process by the actors involved to view value or experience" [3]. In non-sharing services, it is enough to see the value co-creation between a service provider and a consumer (Figure 7).

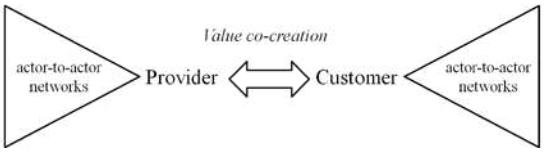


Figure 7. Value co-creation as resource integration process

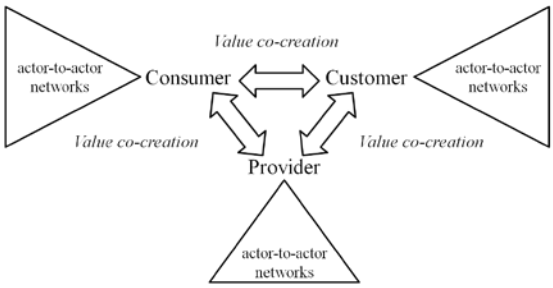


Figure 8. Value co-creation under the supervison of the third actor

However, in sharing services, we need to look at the triadic value co-creations among a service provider and two consumers (Figure 8). The point is that *the value co-creation between two actors is realized under the supervision of the third actor*.

When using Zipcar, one expects that clean cars are always available whenever necessary in exchange for keeping the borrowed car clean and returning the car on time, and this is intentionally achieved with the guidelines and the call centers as described in Section 3.2. In general, the value co-creation between consumers is directly or indirectly protected by the sharing service provider (Figure 9).

In summary, value co-creation that is difficult to realize with just two actors could be realized under the supervision of the third actor that is known to the other two actors. In the terms of our case, value co-creation in sharing services is

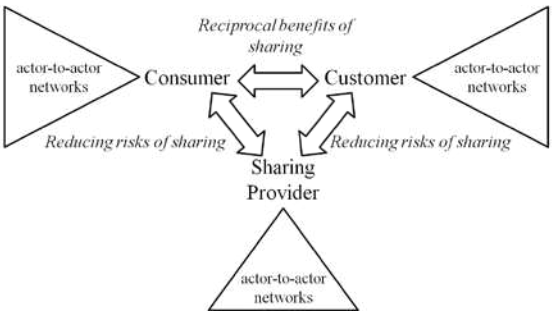


Figure 9. Value co-creation supported by sharing service provider

characterized by triadic value co-creations, where value co-creation between two consumers is realized under the supervision of a sharing service provider.

5 Future work

The 21st century is the service age, and the use of sharing services is a remarkable new trend at the beginning of the century. This paper proposed a hypothesis to answer the question, “Why do people prefer sharing services now?”. Moreover, it restated the hypothesis from the view of service dominant logic, especially in the context of value co-creation.

However, the hypothesis needs to be tested. It could be performed in three steps. First, the analysis of risks and benefits of sharing (as shown in Table 2) will be conducted with as many business cases as possible. Second, to reduce the risks of their services, investigations of the providers’ policies or strategies will be performed. It could be performed through analyzing documents of the services or interviews with the providers. Third, the questionnaires or experiments of the services will be conducted with customers to find out how the they evaluate the policies or strategies and what types of sharing services are really preferred.

References

[1] Botsman, R. and Rogers, R., What's Mine Is Yours: The Rise of Collaborative Consumption, HarperBusiness, 2010.

[2] Wood, J. B., Hewlin, T., and Lah, T., *Consumption Economics: The New Rules of Tech*, Point B Inc, 2011.

[3] Lusch, R. F. & Nambisan, S. (2012). Service Innovation: A Service-Dominant (SD) Logic Perspective. Available at http://air.eller.arizona.edu/docs/papers/2012/Service_Innovation_Lusch_Nambisan.pdf, Last accessed 28 Aug 2013.

[4] Kohda, Y. and Masuda, K., Why do people select the sharing service?, *the First Domestic Events of the Society of Serviceology*, AP-12, Kyoto, 2013.

Service Value Co-creation for Enterprise IT Solution Services

Yukiko Nishioka†, **Michitaka Kosaka**‡

†ACT Consulting Corp.
yukiko_nishioka@act-consulting.co.jp
‡Japan Advanced Institute of Science and Technology
kosa@jaist.ac.jp

Abstract

Service value co-creation is an important issue in service science. We proposed KIKI model as a standard procedure for service value co-creation in B to B collaboration [1]. On the other hand, MUSE (Methodological Universe for the Services Environment) was developed as a new methodology for IT solution services [2]. From the view point of service value co-creation, both have similar characteristics. In this paper, the relationship between KIKI and MUSE are discussed and it is shown that MUSE process can be specified by the standard process for service value co-creation in KIKI model. The value co-creation process using MUSE is introduced through a case study.

Keywords: KIKI model, MUSE, IT solution service, Service value co-creation, Objective value, Functional value, Design office

1 Introduction

In current enterprise IT solution services, many customers cannot utilize IT systems effectively nor reap any benefit from their services due to problems encountered and failures. To rectify this situation, we have analyzed these issues by applying Service Engineering (SE [3])'s viewpoints. Value co-creation and collaboration among customer and IT service providers in the early stage are important to create a basis of the following steps of IT solution services. Furthermore, it shows that significant activities of the design office and effective use of MUSE are valuable and indispensable for achieving the customer's end goal [4-7].

Through those analyses, we introduce the concepts of two IT service values in IT solution services. The first is an "objective value" in grand design phase and the second is a "functional value" in IT service realization phase. Those values are correlated and it is crucial to clarify the relationship between them for successful IT solution services. Collaboration is important for realizing and achieving the values.

In the meanwhile, we proposed KIKI model as a standard procedure for service value co-creation in business to business (B to B) collaboration. Service value co-creation process seems to have similar characteristics in IT solution grand design phase. Through the comparison, we have recognized that MUSE would be a practical methodology for service value co-creation in KIKI model. Thus, KIKI model is very effective for service value co-creation in IT solution service domain as well.

In this paper, first, KIKI model is explained briefly. Then, the method of MUSE and its relationship with KIKI model are described. Next, a case study is reported for demonstrating the effectiveness of service value co-creation for IT solution services facilitated by the design office using MUSE methodology. Finally, we refer to the future of IT solution services.

2 KIKI model for service value co-creation in B to B collaboration

2.1 B to B collaboration model

Customers in a business to consumer (B to C) service are individuals and service values differ depending on each customer's characteristics or tastes. Therefore, service providers introduce personae [8] by defining their life styles and characteristics to design standard services toward the introduced personae. However, service values in B to B service are defined between service provider and receiver as collaborators, especially in B to B IT solution services.

We can see from Fig. 1 both collaborators offer reciprocally necessary information and/or support the actions of their partners. Services are activities to support human beings or organizations to enable them to achieve their objectives or desires. We consider collaboration as a co-creation process where collaborators have service behaviors that benefit each other.

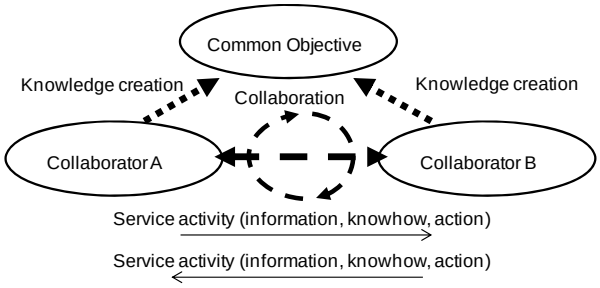


Figure 1 B to B collaboration

2.2 Service value co-creation process in KIKI model

Service value depends on the relationship between the service and its situation, and a more suitable service can be provided if the service field is identified. From such considerations, KIKI model was proposed, where we devised the following four steps for service value co-creation in B to B collaboration [1].

Step.1 (K1). Knowledge sharing related to collaboration: The collaborators in the service value co-creation process understand and share the objectives of the B to B collaboration and its service field, which consists of service providers, service receivers, and the environment around the service. Therefore, the collaborators share knowledge and information related to their purpose.

Step.2 (I1). Identification of the service field: The service field is identified using various technologies such as data mining and questionnaire analysis or collaborations between providers and receivers of the service. What kind of service support is needed for the receivers is investigated.

Step.3 (K2). Knowledge creation for the new service idea: Suitable service behaviors are designed after understanding the service field. Through participants collaborating in the service value co-creation process, new knowledge for service is created by combining various service ideas and technologies.

Step.4 (I4). Implementation of the new service idea: The created new service idea in Step 3 is implemented by considering business model, pricing of services or required IT systems. Collaborators in service value co-creation process evaluate the results of knowledge creation step for the required service and take them into account in the following process for enhancing services.

The above four steps in the service value co-creation process can be described on a two-dimensional plane, as shown in Fig. 2.

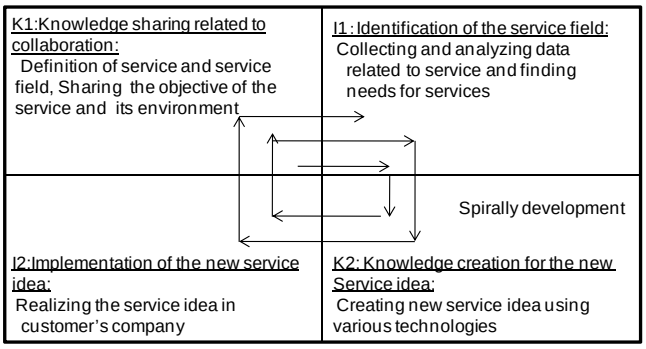


Figure 2 KIKI model

The service that results from this service value creation process can be further enhanced by repeating these four steps of service value co-creation in a spiral of development. As the value co-creation process is repeated, collaborators come to understand the service field much more fully. The result is that the participants eventually understand the needs for services more deeply and generate more suitable service ideas.

3 Service value co-creation in IT solution services

3.1 Two values in IT solution services

There are two IT service values in IT solution services. The first is an "objective value", which is the customer's end goal achieved by utilizing IT solution services. The second is a "functional value", which is a function provided by actual IT system or related services. These two values are intrinsically related and their relationship is shown in Fig. 3 by using SE notations.

There are two phases in IT solution services, which are the grand design phase and the IT ser-

vice realization phase as illustrated in Fig. 3. The upper part is the grand design phase where a customer's “objective value” is extracted and committed to by the customer.

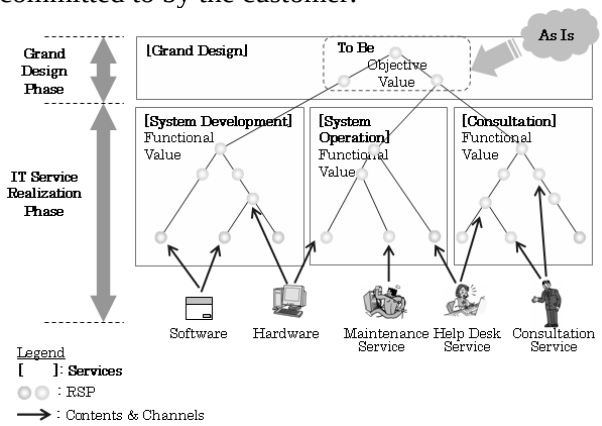


Figure 3 Value chain in IT solution service

The lower part is the IT service realization phase where the “objective value” is translated into “functional values”. These “functional values” are broken down into a detailed functional structure, such as system architecture, subsystem and applications, then screen layout and data to be handled in an IT system. After software, hardware and other related services are developed and integrated into upper level of services in the IT service realization phase, the utilization of services with “functional values” and those interactions can finally achieve the “objective value”.

3.2 Importance of design office

Although the two values are correlated and chained, they still remain separated in many cases. Consultants are usually devoted to extracting the “objective value”, while development companies concentrate on accomplishing the development of IT systems, and operators are mostly concerned with regular operations. These people in these businesses are focusing on realization of “functional values” and care little about whether the “objective value” has been achieved or not, and sometimes neither does the customer who is appropriately satisfied with available services.

It is essential to observe and manage entire IT solution services from a bird’s eye view from a consistent perspective throughout the IT solution service lifecycle. In addition, without any strong determination by executives or collaborative work among stakeholders, these services stagnate and become stacked, and it is hard to ascend to the stage of achieving the “objective value”.

Therefore, the concept of the “design office” is introduced, who is not only responsible for the design of the IT solution services but also for achieving the customer’s “objective value”. Its roles in the grand design phase are value co-creation with the customer and substantiating its knowledge and needs into the grand design of entire services. It promotes stakeholders throughout the service as a facilitator and a producer, also provides business process reengineering (BPR) consultation services, project management services and follow-up services, such as training, helpdesk support in the IT service realization phase.

The name “design office” here represents its role, and it does not matter who plays. The reason we present the design office as a single agent is because we think that its role is very important for overseeing entire services from a consistent perspective throughout the IT solution service lifecycle. Furthermore, an important role for the design office is to manage and adjust all parties, including stakeholders, IT development companies and other IT service providers, which are taking part in pursuing customer levels up to the “objective value”.

4 MUSE Approach

In this section, we introduce MUSE in the grand design phase, which is a practical way of service value co-creation in KIKI model.

4.1 Procedure for grand design phase

The procedure for the grand design phase can be seen in Fig. 4. This phase is divided into six steps. The purpose of steps 1 and 2 is to understand AsIs (e.g., current business situation and current issues) and to share a common perception with stakeholders. We mutually recognize the problems to be solved in these steps, and describe the structure and dynamics of the enterprise as the AsIs model.

Next, ToBe (where we clarify the vision and the end goal of the customer in the future environment) is discussed in step 3. The future vision and the end goal are clarified in step 4, where the

view of the structure as well as dynamics in the future is modeled and drawn in a picture. The "objective value" of IT solution services becomes obvious and specific through these steps.

Then, the "objective value" is translated into "functional values" in step 5, where the entire IT solution services are overviewed from a bird's eye view. Finally, step 6 involves planning and budgeting of the discussed grand design.

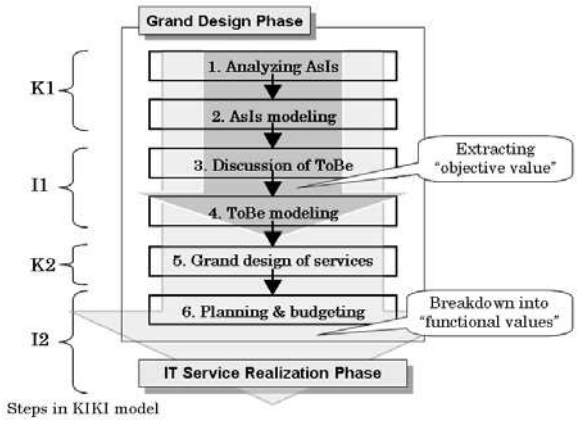


Figure 4 Procedure for grand design process

4.2 MUSE method as a communication tool

MUSE method plays two roles. The first is as a communication tool and the second is as a modeling tool.

MUSE method as a communication tool is used in steps 1 and 3. In usual case, several teams are formed, such as workers who present work tasks in the enterprise business cycle, middle managers and executives. AsIs problems and the ToBe end goal were repeatedly discussed in a brainstorming style by each team. In MUSE, we have merged brainstorming [9] and KJ-method [10] and expanded into the anonymous way of discussion. It consists of the following six steps.

(1) Participants are seated around a table.

(2) Participants write down their opinions on post-it notes according to the theme of the session. All notes are gathered and shuffled, and distributed back to each participant.

(3) A session consists of the following steps.

(a) First, a chairperson selected by the participants chooses one of the allocated notes. The chairperson reads aloud the opinion written on the note, and places it on the MUSE sheet.

(b) Then, the chairperson forces other participants who have the same or similar opinions on their notes to read the opinions on them and express the points of similarity. The chairperson has the rights to accept or reject the proposals of participants.

(c) In this way, all notes with the same or similar opinions are presented and posted near the chairperson's note on the sheet, and then the first session is over.

(d) In the next session, another participant becomes the chairperson, and the session then starts.

(4) Sessions are continued until all notes are posted on the sheet. In this way, the participants' opinions are gathered and categorized in the form of islands of notes with the same or similar opinions. Fig. 5 is a snapshot of a session using MUSE method.

(5) After the above sessions, an appropriate title for each category is discussed and written on the label, different color note, and placed on top of each category.

(6) The relation map of title labels is configured in tree style according to the important axes such as cause and effect, or principles and practice.

Figure 5 Snapshot of a session using MUSE method

MUSE has four main characteristics: (1) democratic, (2) game sense, (3) rapid analysis, and (4) breakthrough, which are due to the following facts.

The anonymous way leads discussion in democratic manner, without any constraints of age, sex, experience, power, or raised voices, and participants can think deeply and understand others' opinions that they share until the discussion ends. Participants also enjoy the discussion in a game sense to win the debate and finish up the allocated notes as earlier than anyone. Furthermore, even if the participants have never met before, the total time for all sessions usually only takes two to three hours, and importance and

concerns felt by the participants can be over-viewed on the relation map. In addition the consciousness level of participants is improved through intensive discussions among participants from different points of view, and as a result unexpected breakthroughs are often achieved.

4.3 MUSE method as a modeling tool

MUSE method is used as a modeling tool in steps 2 and 4 in the grand design phase. While in step1 or 3, discussions among the participants are emphasized, in step2 and 4, over-viewing the enterprise activities are emphasized. Problems and values are extracted from the overall viewpoint by walkthroughs in the AsIs and ToBe models.

The AsIs modeling of MUSE method consists of the following steps.

(1) Extract data sets from existing documents which are related to the customer's business activities.

(2) Find agents who use the above data sets.

(3) Extract the functions of the agents, and describe them as the actions of the agents.

(4) Find artifacts associated with the business, such as facilities, manufacturers, end customers, and so on.

(5) Describe those agents, data sets, and artifacts in post-it cards, using MUSE notation. Fig. 6 shows the MUSE notations for modeling.

(6) Draw the overall image of work tasks, through overlooking a set of agents and data on the MUSE sheet.

(7) Place artifacts at the proper position and mark boundary lines on the MUSE sheet for making clear the region of each section. Fig. 7 is an example of MUSE modeling (AsIs).

In this AsIs modeling, work tasks and work flow dynamics are verified afterward by the walkthrough procedure together with the customer and the design office, which means to walk virtually along work flows in the model. Problems such as loss, complexity, inconsistency, overlaps of work tasks are investigated through such over-viewing and walkthrough processes.

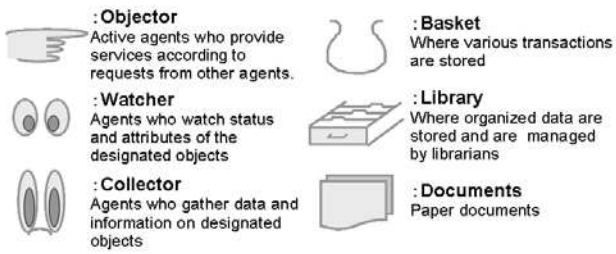


Figure 6 MUSE notations for modeling

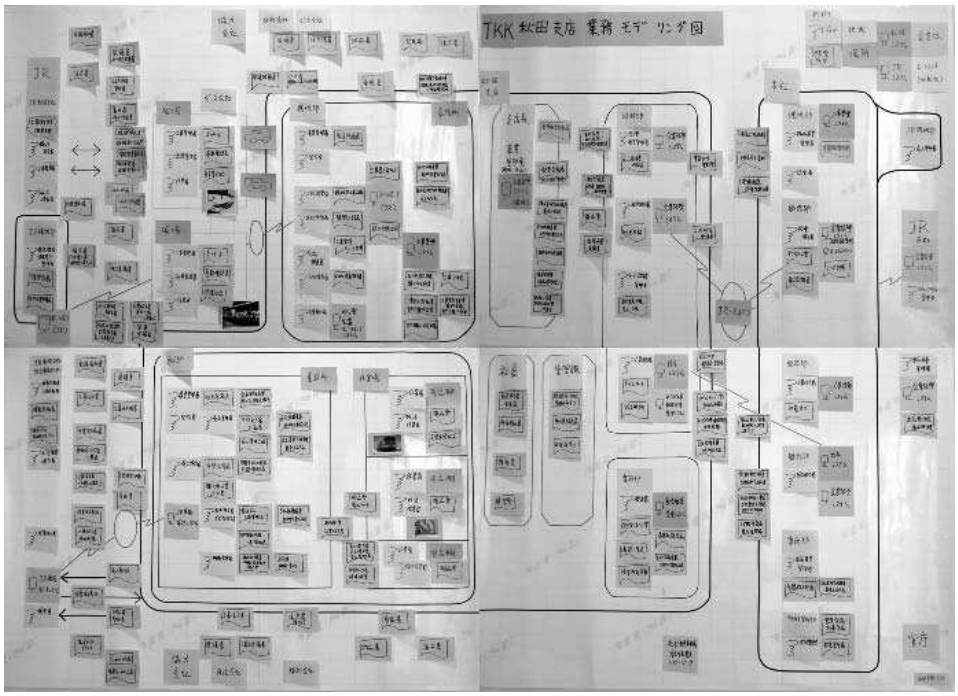


Figure 7 An example of MUSE modeling (AsIs)

Although AsIs is vague and unclear at the beginning, it is important to identify inconsistent and conflicting current situations, which are mirrored explicitly in the overall picture on the MUSE sheet.

As the next step, we draw the ToBe model. First, we analyze and decompose the work tasks and into function level. Then, we recompose ToBe functions and agents who serve those functions. For drawing ToBe model, we arrange agents in the optimal positions with proper data in the MUSE sheet. Following three aspects are important points to determine where to place the agents on the sheet: (1) to follow up the vision of the customer, (2) to be free from the current rules and organization, and the conventional manner, constraints and customs, (3) to define the measures for evaluation of the overall system.

The structure and dynamics of the future organization described in ToBe model is surely achieved by the capabilities of IT and BPR. IT significantly contributes to overcome time and space differences, and volume and accuracy constraints, and BPR also contributes very much to remove organizational, rules, and other environmental constraints.

The characteristics of MUSE modeling are: (1) overall enterprise activities are visualized, (2) properly analyzed, and (3) the structure and dynamics are overviewed within a short time by a small group of people.

Those characteristics are resulted from the uniqueness of MUSE modeling method, such that (1) there are no process flows in the MUSE modeling sheet, (2) because iconic notations of agents and data are used to illustrate business activities within the company and inter companies, it is easy to understand by those who are not familiar with IT terminologies, (3) no-one could draw AsIs or ToBe modeling picture in free hands, but you can draw them according to MUSE modeling steps, and (4) walkthrough in a model helps us to verify the work tasks and work flow dynamics from various view points, such as a certain agent's view point or a certain data's lifecycle, and problems and values rise to the surface.

4.4 MUSE as a practical methodology for KIKI model

Using KIKI's terminology, the purpose of MUSE in grand design phase is to share the service system itself and to build service field for value co-creation within the stakeholders. Step1 and 2 correspond to K1, step3 and 4 correspond to I1. Because the clarification of "objective value" corresponds to I1 (Identification of the service field associated with customer's requirements) and the "functional values" to K2 (Knowledge creation for the new service idea). Therefore, Step 5 corresponds to K2 and Step 6 and the following IT service realization phase to I2.

Also, we claim that the role of the design office, which knows both customer and IT vendors, is important for such collaborative works. Thus, MUSE steps conducted by the design office can be considered as a practical methodology for KIKI model in IT solution services.

5 Case Study

5.1 IT solution service for a utility company

As a case study of IT solution service, we discuss the case of a utility company. The IT system for the facility management division of the company was developed and completed in a short time and has been successfully operating since 2003 [11]. The objective of the IT system is to realize PDCA (Plan, Do, Check and Action) cycle for managing lifecycle of facilities, through reengineering and improving work processes. We have engaged in this IT solution service as a design office.

This IT system is used by more than 2,000 workers of the company, and has been continuously enhanced since the beginning of operation. Because of such successful deployment in the above division, the IT system has been extended in large to other 3 divisions and fully contributed to improvement and restructuring of the business activities of the company. As a result, the whole system saved financially one half of the investment cost as compared with that of traditional approach. Due to such a meaningful contribution, this project won the 2005 Shibusawa Award and the 2006 Ohm Award (prestigious Japanese awards in the filed).

5.2 Value co-creation among the stakeholders in grand design phase

The grand design phase has started from the investigation of the customer's activities. In the above case, there are various stakeholders as

follows. There are the head office, branches, and field maintenance offices inside the division, other divisions in the company, contractors outside the company, and end customers who finally receive utility services.

In the process of analyzing AsIs of activities in the facility management division, we, as a design office, realized that so many tasks were overlapped and complexly related, and their definition and responsibility were vague, and furthermore rules among them were often unwritten. There was no one who was able to describe the overall activities of managing facilities. Then we thought that our developed methodology MUSE would be useful to overview those work tasks and their complex relationships. From this view point, we propose MUSE and the customer accepted it.

The procedure for grand design phase shown in Fig. 4 was executed using MUSE communication tool and modeling tool. After current activities analysis in step1 and 2 stakeholders understand and share the real problems of AsIs. Those steps correspond to K1 in KIKI model.

In step3, the vision of the facility management division and its core concept were discussed. Through this discussion and walkthrough in the ToBe model in step4, the vision and concept were redefined. A catchphrase, "From Construction to Maintenance" was adopted. This means that the period of high economic growth has been ceased, and we should change the facility management style. That is, it is more important to focus on maintaining currently available facilities rather than to construct new facilities. Those steps correspond to I1.

According to this concept, the "objective value" of the IT solution service was discussed and specified as "To realize PDCA cycle for managing lifecycle of facilities through IT services". To satisfy this objective, the requirements of IT system and related services with the "functional values" were discussed in step5 (K2).

Then those new IT services were planned and budgeted in step 6 and were realized in the following phase described in 5.3 (I2).

According to those steps in MUSE methodology with the design office support, we could reveal and clarify the customer's "objective value" in a short period and go forward to IT service realizing phase. Although it became harder than before for customers to optimize and balance the overall system goals by themselves, collaborative work was a crucial factor to achieving effective services. As is shown in this case study, MUSE and the role of the design office are practical and reasonable way for service value co-creation in KIKI model.

5.3 Value co-creation among customer and IT venders in IT service realization phase

Following to the grand design phase, IT service realization phase started, which is a later part of I2 in KIKI model. In this phase, the collaborative work among customer and IT venders, especially IT users in the customer and system engineers of IT development companies, became important.

In this case, the "objective value" was "To realize PDCA cycle for managing lifecycle of facilities through IT services". There is dilemma that users of IT systems could not assess how to improve their business process due to inadequate knowledge about IT technologies, and IT system development companies could not effectively utilize IT technologies due to insufficient knowledge about customers' business processes.

This dilemma was solved by organizing several joint teams of IT users who represented work tasks in the facility division and system engineers at IT system development companies. The members of those teams exchanged ideas and had discussions toward accomplishing the user's final objective. The ideas of managing lifecycle of facilities using an IT system were discussed and reviewed repeatedly by the project managers of the customer. This process was continued and repeated over short periods and "functional values" are broken down into details.

Through those value implementation processes, the IT service realization phase has successfully implemented and the main subsystems of IT system for facility management were developed and started operation within 18 months. As a result, the system could change customer's work styles into a flat organization without middle management and resolve sectionalism issues by using the same information and making their business activities more efficient. The principle behind IT solution services could successfully be changed "From Construction to Maintenance".

6. Conclusion

We have introduced by showing successful case

study: (1) there are two values i.e. "objective value" and "functional values" in IT solution services with value chain, (2) significant activities of the design office and effective use of MUSE are valuable and indispensable for clarifying and achieving the customer's end goal, which is the "objective value", (3) collaborative work among customer and IT venders becomes a crucial factor to realizing effective services, which are "functional values". Those are the procedure of K1, I1, K2 and I2 in KIKI model and MUSE methodology and the role of design office are the practical way for actuating KIKI model in IT solution services.

In the last, we would like to mention that IT solution services have been playing important roles in achieving a future vision for customers. We need to manage IT solution services sustainably to effectively and continuously utilize IT systems by planning IT strategies, drawing up roadmaps to achieve the strategies, executing the roadmaps, and modifying the strategies depending on the changing environment. The design office, from this viewpoint, should clarify the relationship between business strategies and IT solution services to explain the necessity for IT systems. Collaboration and service value co-creation among stakeholders are necessary to plan such IT strategies and effectively execute IT solution services. The role of the design office, which maintains a consistent view of IT solution services and prepares opportunities for collaborative discussion, is crucial in providing successful and sustainable IT solution services.

However, the most important success factor is the customer's will to effectively utilize IT solution services in their business. The design office and MUSE will support the customer's resolve as far as there is a strong will to attain IT values in enterprise by co-creation among related parties.

Acknowlegement

This research was supported by Japanese Grants-in-Aid for Scientific Research (KAKENHI No.24510186).

References

[1] M. Kosaka, Q. Zhang, W. Dong and J. Wang, Service value co-creation model considering experience based on service field concept, *Proc. of IEEE International conference on service system and service management*, pp.724-729, Shanghai, China 2012.

[2] Y. Nishioka, Effectiveness of MUSE methodology in grand design process of IT system construction projects, *The 30th Workshop on IEEJ-Information Systems*, 2007 (in Japanese).

[3] Y. Shimomura, T. Sakao, L. Petti and A. Raggi, Proposal of a service design process model based on Service Engineering, *Proceeding of the 6th International Symposium on Tools and Methods of Competitive Engineering*, Ljubljana, Slovenia, CD-ROM, pp.1665-1675, 2006.

[4] Y. Nishioka and N. Koike, A prospect on IT business from a view point of the Service Engineering, *Journal of the 16th System Design and Dynamics*, JSME, No.06-33, pp.84-87, 2006 (in Japanese).

[5] Y, Nishioka and K. Yamamura, Case study (4) IT solution service; In Tokyo University Ed., Reports of engineering methodology for service expression and evaluation aimed at customer satisfaction and productivity improvement, *MITI undertaking investigation for service productivity improvement*, pp.113-131, 2008 (in Japanese).

[6] Y. Nishioka, Realizing the objective values of customer in enterprise IT solution services -Service oriented requirement development method "MUSE"-, *IEEE International conference on service system and service management*, CD-ROM, SP050, Tokyo, Japan 2010.

[7] Y. Nishioka, Achieving objective values led by the "design office" - A study of success factors in the IT solution services which solved business problems -, *Journal of Digital Practice*, vol.1 No.4, pp.190-199, 2010 (in Japanese).

[8] A. Cooper, (1995). About Face: The Essentials of User Interface Design, John Wiley & Sons, 1995.

[9] A. F. Osborn, *Applied imagination, 3rd revised edition*, Charles Scribner's Son, 1963.

[10] J. Kawakita, Hasso-ho (abduction method), Chuokoronsha, 1967 (in Japanese).

[11] T. Uemura, T. Oiki, M. Oka, and Y. Nishioka, Kyushu Electric's ERP system development and operation for overhead lines system, *Conseil International des Grands Réseaux Électriques (CIGRE)*, CD-ROM, B2/D2-105, pp.1-8, 2006.

A Context-aware Information in Smart Home for Health Recommendation Service based on CARE Architecture

Konlakorn Wongpatikaseree **Azman Osman Lim** **Yasuo Tan**
School of Information Science
Japan Advanced Institute of Science and Technology, Ishikawa, Japan 923-1211
{w.konlakorn, aolim, ytan}@jaist.ac.jp

Abstract

Nowadays, providing a health service to home users presents a number of challenging tasks. Several researches attempt to provide health services to home users. Apart of health service, health recommendation service is proposed as a term of recommendation health service for preventing the disease. However, the existing research works suffer to provide the appropriate health recommendation service because most of researches analyze the disease based on only user's health condition. In this paper, we aim to help the physician to analyze the health recommendation service by providing the human behavior information and context-aware information in smart home. The healthcare system and physician can analyze the cause of disease through provided information. To obtain there relevant information, context-aware activity recognition engine (CARE) architecture is proposed for providing the context-aware information and also human activity information. Human behavior analysis system is also established in this research for summarization the outputs from the CARE architecture. Through experiment studies, the results of questionnaire express the importance of the provided information. The human behavior information can be used to analyze the appropriate health recommendation service for the home user.

Keywords: Healthcare system, Health recommendation service, Human behavior, Context-aware information, CARE architecture

1 Introduction

In the 21 century, concern about healthcare has become an essential aspect in daily life, not only just for the elderly but also for young people. People are more concerned about their health and desire healthy. Recently, healthcare systems have become more popular, especially in the smart home domain. The smart home [1] has emerged as one of the mainstream approaches to support technology-driven independent living for elderly and disabled persons. The smart home concept is gathered with several technologies i.e. wireless sensor network (WSN), data communication, and security to produce the ambient intelligence in home. Hence, several healthcare systems have proposed belonging to the smart home concept, e.g. a home healthcare (HHC) system [2].

According to the HHC system, there are many companies have released the products, which use for monitoring daily health conditions. Panasonic has released EW-NK63 device for measuring the calories consumed in each day. OMRON has presented HeartScan device for monitoring and recording the ECG information. Based on current technologies, users can maintain the health condition by themselves, and physician can utilize these data for diagnosis the diseases. Nonetheless, only basic health condition might not enough for analyzing the patient's health in some cases. It is difficult to accurately assess an individual's health condition because each person has a different lifestyle. Moreover, the existing healthcare system cannot find the real cause of the disease because only health signals are used to diagnose the disease. The cause of the disease does not directly appear in the health signals. Therefore, there are several types of data available which have to be considered such as user location, or user action. Consequently, human behavior information plays an important role to help the healthcare system to diagnose and prognoses the disease.

To address this shortcoming, E-health is proposed to manage the user's health information. It removes the need to process paper-based, and allows healthcare system or physician to record the the user's health condition electronically. Thus, in this paper, we aim to provide the human be-

havior information to help the physician to analyze the health recommendation service for the home user. Consequently, context-aware activity recognition engine (CARE) architecture is introduced for obtaining the human behavior and context-aware information in smart home. Human behavior analysis system, demonstrated as the E-health, is also implemented on top of the CARE architecture for summarization the human behavior information. The results of this system can aid in physician diagnoses, enabling them to make more accurate recommendation on how to prevent disease. Through this studies, the questionnaire's results express the effectiveness of our proposed system.

2 CARE Architecture

To carry out the goal of this paper, CARE architecture is designed as human activity framework. The application that requires the human activity information can build on top of the CARE architecture. The proposed CARE architecture has principal task to improve the ability of activity recognition system and provide the results that are more accurate, more reasonable, and more reliable in the real environment. Several techniques are applied in the CARE architecture. For example, BSN and HSN techniques are developed to collect the human information and surrounding information in smart home. To classify the human activity, this proposed architecture focuses on 14 target activities that users often perform at home, such as "Watching TV," "Working on a computer," or "Cooking" activities. The CARE architecture consists of six layers, which combine several technologies and techniques inside as shown in Figure 1.

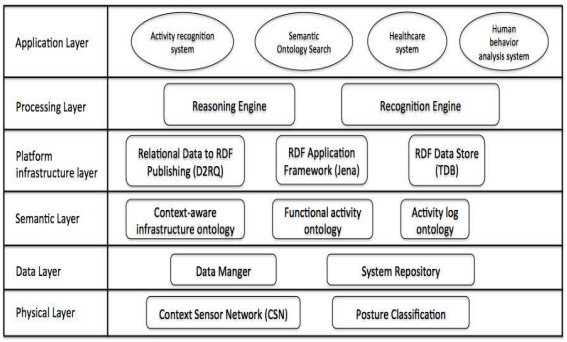


Figure 1. A layered architecture of the CARE architecture

2.1 Physical Layer

Physical layer consists of physical hardware such as sensors, home appliances, or network components. This layer provides the context-aware information in smart home, including the human information. To obtain the relevant information in a real environment, Context Sensor Network (CSN) with a diversity of sensors and network protocols is proposed. The CSN is a sensor network that typically collects the context-aware data in the smart home. There are three kinds of sensor networks in the CSN: a home appliance sensor network, a home furniture sensor network, and a human sensor network. The goal of first two sensor networks is to capture the object usage in smart home. The variety of sensors are built-in the home facilities such as "Sofa," "Outlet," or "TV". Whereas the human sensor network is proposed for observing the human information, such as human location in smart home. Moreover, several communication networks are also developed in these three sensor networks: UPnP, ECHONET, and Zigbee.

Nevertheless, using information concerning object activation and human location still has limitations because sometimes, human location does not hint at any specific activity if several objects are being used. In this layer, posture classification is proposed to improve the ability of activity recognition. A novel algorithm, called range-based algorithm [3], is used for classifying the human posture. The concept of this algorithm is to use the relation of body parts to extract the postures from the range between body parts.

2.2 Data Layer

Organizing the huge amount of data in a smart home is a complicated task because there are several information components for the system to consider. The system cannot always obtain absolutely perfect data from the sensor. There can be missing data and noise problems associated with hardware. Therefore, the data layer has the responsibility of storing and organizing the data obtained from the physical layer. A data manager and system repository are proposed to organize the data in the smart home. The data manager is proposed to normalize and transform the data before sending it to the system repository. There are two main functions to normalize raw data: supply missing data function and eliminate

data function.

Meanwhile, the system repository plays a vital role in controlling data flow in the system. There are two modules in the system repository: a repository and a context controller. The repository is used as the database in the CARE architecture. It collects the data that is normalized by the data manager and also retains the temporal reasoning of the system. The context controller performs three main data-processing tasks: "Mapping Data," "Composing Data," and "Reprocessing Data". This context controller will synchronize the data between data layer and semantic layer.

2.3 Semantic Layer

In semantic layer, three ontologies namely context-aware infrastructure ontology, functional activity ontology, and activity log ontology, are designed for creating smart home knowledge-based for recognizing human activities.

The proposed semantic layer makes CARE architecture scalable because it describes at the abstract level, so it does not need a training process. It differs from the existing researches, which use the probabilistic method to recognize the human activities. They require a large amount of data for training set. Furthermore, it also gains the benefit from the semantic web technology – standards for distributed data sharing and processing. Based on these three ontology models, two pieces of information are introduced for improving the ability of activity recognition: the new user's context and Activity's Location in the Activity Log (AL^2). For the new user's context, we add the human posture information, which is obtained from the posture classification, to reduce the "ambiguous activity problem". Meanwhile, the AL^2 is introduced to find the relationship between activities that occur at the same location. The history of activities at the current user's location is investigated with the current activity for classification [4].

2.4 Platform Infrastructure Layer

To translate from the data to the knowledge, ontology application manager management (OAM) framework [5] is developed. The OAM framework has a duty to link between a unified conceptual backbone for data modeling and representation and an application layer. Resource description framework (RDF), a standard model for data interchange on the web, is created via the OAM framework. The RDF will collect the data modeling, which is mapped between the data in the repository and the ontology model. After that, all of data modeling will be stored in the smart home knowledge-based. Moreover, the OAM framework also provide the engine to access the knowledge through SPARQL. The SPARQL is a language query that used to query the knowledge from the RDF (refers to the ontology model and database) in a smart home knowledge-based.

2.5 Processing Layer

The intelligent technique is built in the processing layer for recognizing human activity based on the smart home knowledge-based, which is created in the platform infrastructure layer. Normally, ontological models do not have the ability to recognize human activity. The popular ontologial language, OWL (Web Ontology Language), has been used to build activity ontologies, and to recognize activities based on context data. Naturally, ontology uses Description Logic (DL) to express domain knowledge. To infer the human activity, DL rules are created by modeling and linking between an activity instances and inferred activity instance. DL rules in this research describe relationships among object activation, human location, human posture, and activity log. The following example indicates a DL rule for the "Washing dishes" activity:

$$
\begin{aligned}
&Washing\ dishes \sqsubseteq Functional\ Activity \\
&\sqsubseteq Kitchen\ Activity \\
&\sqcap\ use(Object.Furniture(Sink)) \\
&\sqcap\ Object.Human.Current_location(kitchen) \\
&\sqcap\ HumanPosture(Stand) \\
&\sqcap LastActivity.KitchenActivity(Eating\ or\ drinking)
\end{aligned}
$$

2.6 Application Layer

For the application layer, several applications, which require the human activity information and semantic context-aware information in smart home, can be implemented on top of the CARE architecture, such as activity recognition system, healthcare system, or human behavior analysis system. To evaluate the performance of activity recognition based on the CARE architecture, the proposed architecture is installed in iHouse [6],

an experimental smart home environment. The experiment was performed by six actors (three males and three females) whose ages ranged from 24 to 31. They were asked to perform any activities in iHouse without instruction. The results of the activity recognition based on the CARE architecture are shown in Table 1.

Table 1. Accuracy of activity recognition based on CARE architecture

Activity	Accuracy (%)	Other possible resultant activities
A_1 = Sitting on the toilet	93.75	A_9 (2.08%), A_{14} (4.17%)
A_2 = Taking a bath	100	—
A_3 = Lying down & relaxing	100	—
A_4 = Sleeping	91.86	A_{11} (3.49%), A_{14} (4.65%)
A_5 = Making coffee	100	—
A_6 = Cooking	86.21	A_{14} (13.79%)
A_7 = Eating or drinking	100	—
A_8 = Washing dishes	100	—
A_9 = Working on a computer	92.25	A_{10} (6.90%), A_2 (0.42%), A_{14} (0.43%)
A_{10} = Watching TV	97.94	A_{14} (2.06%)
A_{11} = Reading a book	100	—
A_{12} = Scrubbing the floor	93.75	A_{14} (6.25%)
A_{13} = Sweeping the floor	96.67	A_{14} (3.33%)
A_{14} = Others	100	—
Average accuracy: 96.60 %		

Based on the goal of this research, we are going to use the human activity information, which obtain from the CARE architecutre, by implementing the human behavior analysis system for analyzing the disease.

3 Actualization of a Health Recommendation Service

Until now, providing the appropriate health recommendation services in home is a challenging task. The system has to decide which health services are suitable for the home user. There are several proposed systems that attempt to provide the health recommendation service to the home user. Nonetheless, most of the existing researches tend to collect only user health condition through intelligent wearable devices based on various kind of techniques, and diagnose the disease based on only the user health condition. SapoFitness [7] has been established to monitor the health information, especially user's weight information, to control the user's weight. For monitoring the heart rate, mobile health monitoring system (MHMS) architecture has been designed to monitor the heart rate from RFID ring-type pulse sensor and display the heart rate through smart phone [8].

Nevertheless, existing healthcare systems cannot find the real cause of the disease because they use only health signals, while other information must be also taken into account, notably human behavior information. Hence, the outputs of the CARE architecture will be used for analyzing and presenting to the home user and the physician for diagnosis the disease through the human behavior analysis system. Both activity information and context-aware information in the home are presented in the human behavior analysis system. Four kinds of information are analyzed in this system. Figure 2 illustrates the example interface of activity information in the human behavior analysis system.

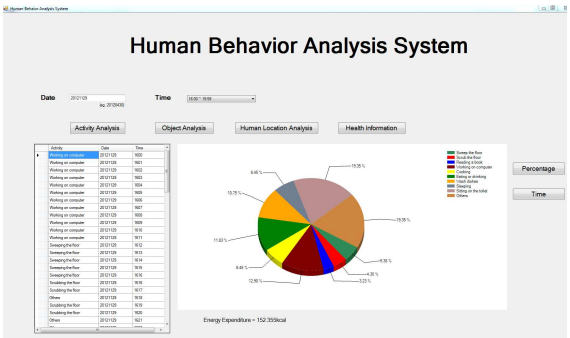


Figure 2. The activity information in the human behavior analysis system

3.1 Activity Analysis

Normally, to diagnose the disease, the health information is the main information that the physician or the healthcare system uses. However, the real cause of the disease may not show off in the health information. Therefore, the activity information plays an important role in recognizing the cause. Based on the summarization of the activity information in this system, three types of information are provided. Activity of daily living (ADL) information is presented to the home user. The home user, physician, or healthcare system can monitor the ADL information to perceive what a home users does in each day for analyzing the cause of disease. Moreover, the activity information is also shown in term of percentage and time.

These kinds of information are very useful for ascertaining the cause of disease. For example, the existing healthcare systems cannot diagnose "Diarrhea disease" because the symptom of "Diarrhea disease" does not appear in the basic health information. Nevertheless, the human behavior analysis system can predict this illness from the "Sitting on the toilet" activity. If the

home user tends to perform the "Sitting on the toilet" often, the healthcare system might predict that the user has some problems with "Diarrhea disease". Consequently, the system can provide a health recommendation service to the user for checking and preventing the "Diarrhea disease".

Consider to the famous disease as "Obesity", it is considered a serious disease since it is one of the main causes of diseases such as "Diabetes", "Heart disease", and "Cancer". Of course, it is easily to diagnose "Obesity" from body weight, but the reason for "Obesity" is not dependent only on body weight. The daily habit of the user can also affect "Obesity" problem. In this research, the human behavior analysis system also provides energy expenditure information based on the activities that user performed in each day. Energy expenditure is calculated based on Metabolic Equivalents (METS) values. The METS value is most frequently used in calorie count to compute energy consumed during each activity. The formula for energy expenditure in this research is defined by the American College of Sports Medicine (ACSM), illustrated in equation below.

$$Energy\ Expenditure\ (kcal) = 1.05 \times METS \times Duration\ (hour) \times Weight\ (kg)$$

Based on the provided energy expenditure, the system can perceive how many calorie burn in one day based on the activities that user performed. Thus, if home user tends to perform low calories activities often, system might serve the health service, such as exercise service to the home user for preventing the "Obesity" problem.

3.2 Object Interaction Analysis

Normally, the usage of objects information is used to recognize the human activity. However, in this research, the human behavior analysis system provides the object interaction information for analyzing in the healthcare domain. The object interaction information can be analyzed in several directions. The system can perceive how long user spend for sleep from detecting the duration of "Bed" usage. The sleeping time is important to know that home user has adequate sleep or not. Monitoring the use of "Sugar container" can help the home user, who has "Diabetes" problem, control the user's blood sugar.

However, analyzing only the object interaction information for healthcare domain might not enough because some objects in home, user turns on the home appliance, but does not use. For example, user might turn on "TV" in the living room while cooking in the kitchen. Thus, the object interaction information might useful when combining with other information such as activity information, or location information. Normally, the object interaction information is linked with the activity information because the object interaction information can be used for classifying the human activity. Thus, combination between the object interaction information and activity information will improve the ability of analyzing the disease.

3.3 Location Analysis

The location information is also provided in the human behavior analysis system. The location information is very useful for a patient who lives alone in home and has Alzheimer's disease. Sometimes there can be a situation where ther person has regular lapses in memory and forgets what/where the activity is that they were supposed to be doing. Caring for a person with Alzheimer's at home can be difficult. The location information can help people in the early Alzheimer's disease. The system can give the health service as "Quiz exercise" for asking when the patient moves from one place to another. For instance, the system might ask or remind the patient for checking the gas system when patient leaves the kitchen.

3.4 Health Information Analysis

Based on the existing healthcare system, the health information is the most relevant to diagnose the disease. Thus, several companies, e.g. Panasonic, OMRON, or Hitachi, have released products for obtaining the health information in home, and also provide the software to monitor the daily health information.

In the same direction, the human behavior analysis system also provides the health information to the users. For example, the weight information is summarized and displayed in line graph for helping the home user perceives the change of weight. Blood pressure information is the basic information that is used to diagnose "Hypertension" disease. Threshold value can be set for detecting the abnormal of the blood pressure information. Based on this graph, the sys-

tem can detect the tendency of the health information. Thus, this system can provide the health service if the abnormal information is detected.

4 Questionnaire

To evaluate this research, questionnaire technique is chosen to decide the performance of the proposed researches. Questions in questionnaire cover all of proposed, CARE architecture, activity recognition, and human behavior analysis system. The questionnaire about the effectiveness of this research is answered by six actors who demonstrate all experiments in this research. The questionnaire consists of three parts, which are following.

4.1 User's Opinion in Information in Healthcare System

First part of this questionnaire asks the user's opinion in information of existing healthcare system and information in our research. Questions in the first part were measured under each statement on Yes-no question and 5-point Likert scales. Mean and standard deviation (SD) values of the responses were calculated according to the 1-5 scores in difference indicator. Table 2 summarizes the response of three questions in the first part of a questionnaire.

The results from Table 2 illustrates that using only the basic health information might not enough. The respondents agreed with the first statement (with a mean score of 2.67) that the basic health information is insufficient for diagnosis the disease. A mean score is lower than half of level Likert scales (3). Therefore, using the basic health information in the healthcare system might not enough for diagnosis in some circumstances. Although, the basic health information can be used for finding the end result, it cannot perceive the cause of disease.

According to the purpose of this research, the knowledge regarding the human activity will be used for diagnosis the disease. The question of the usefulness of the activity information is asked the respondents. The respondents agreed with the statement (with a mean score of 4.67) that the knowledge regarding to the activity information is useful for providing the health recommendation services. The activity information can be used for diagnosis and prognosis the disease.

Meanwhile, the last question in this part mentions about the usefulness of the human behavior summarization. All respondents gave the positive answer with 100% agreed that the human behavior summarization is useful for the home user to recognize what user does in each day. Moreover, this information is also useful when using for analyzing the cause of the disease.

4.2 User's Opinion in Relationship between Providing Information and Diseases

For the second part of questionnaire focuses on the relationship between the proposed context-aware information, which consists of three kinds of information (activity information, object interaction information, and human location information), and the diseases. The question asks the respondents about " Please answer in which you believe that the X_i can be used to diagnose, treat, or prevent the following diseases." where X_i denotes the activity information, object interaction information, and human location. Ten diseases, which are common happened in Thailand, are observed in this research by asked the relationship between these ten diseases and the context-aware information. The questions are measured based on the Yes-no question. In addition, the results were calculated in percentage.

4.2.1 Recognize Diseases base on Activity Information

From the results in Table 3, we can conclude that all respondents believe that the activity information can be used to diagnose, treat, or prevent the given diseases. For example, recommendation the health service for the "Cystitis" and "Diarrhea" diseases is not an easy task because the healthcare system cannot get the health information that relates to the "Cystitis" and "Diarrhea" diseases. However, the proposed system can predict the "Cystitis" and "Diarrhea" diseases if system detects the home user tends to perform "Sitting on the toilet" activity often.

In addition, the ADL information in this system also helps the healthcare system not only to provide the appropriate health recommendation service, but also to analyze the cause of disease. The result in "Obesity" is a good outstanding example that activity information is useful for the healthcare system. Since the main cause of "Obesity" is that the home user has physical inactivity, so the physical activity on free time are

Table 2. User's opinion in Information in Healthcare System

Statement	Indicator	Min	Max	Mean	Standard deviation (SD)
Do you think the basic health information such as weight, blood pressure, or step count is enough for diagnosis, treatment, and prevention of diseases?	1 (Insufficient) - 5 (Sufficient)	2	5	2.67	1.105
Do you think that knowledge regarding the human activity is useful for providing health recommendation services?	1 (Useless) - 5 (Useful)	4	5	4.67	0.471
Is Human behavior summarization useful for the home user?	0 (No) , 5 (Yes)	5	5	5	0

Table 3. Summarization of the relationship between three kinds of information and the diseases

Diseases	Activity Information		Object Interaction Information		Human Location Information	
	Yes	No	Yes	No	Yes	No
Migraine	50 %	50 %	67 %	33 %	33 %	67 %
Tension-type headache	83 %	17 %	100 %	0 %	50 %	50 %
Cystitis	83 %	17 %	50 %	50 %	57 %	33 %
Chronic Fatigue Syndrome	50 %	50 %	50 %	50 %	67 %	33 %
Hypertension	50 %	50 %	17 %	83 %	33 %	57 %
Obesity	100 %	0 %	50 %	50 %	67 %	33
Diabetes mellitus	50 %	50 %	33 %	67 %	17 %	83 %
Hypercholesterolemia	50 %	50 %	33 %	67 %	0 %	100 %
Diarrhea	83 %	17 %	83 %	17 %	100 %	0 %
Alzheimer	67 %	33 %	50 %	50 %	17 %	83 %

insufficient or too irregular to control the target body weight. In this sense, when the healthcare system detects that the home user tends to perform physical inactivity The healthcare system might serve the health recommendation service as "Body-weight control service" to the user.

4.2.2 Recognize Diseases based on Object Interaction Information

From the results, the respondents believe that the object interaction information is not important for diagnosis the diseases. For example, there is no home facility that can infer to the "Hypercholesterolemia", or the "Hypertension" disease cannot detect from the home facility usage. Nevertheless, not all of home facility do not have a link to the diseases. There are some diseases the respondents believe that it might infer to the diseases. For instance, using "Lavatory" often can infer that the user might have some "Diarrhea" problem, or "Tension-type headache" disease might happen because home user uses the "Computer" for a long time.

4.2.3 Recognize Diseases based on Human Location

The average results of human location got the lowest score when compare with the activity information and the object interaction information. The main reason is the human location does not hint any information that related to the diseases. It means that to diagnose the disease, it does not need the human location as the main information for analyzing the disease. However, consider in detail of disease, the respondents believe that there is a location that can infer to disease, especially "Toilet". Absolutely, the human location might infer the disease, but it does not do the important information when diagnose the disease. It might be worked when using with other information such as the activity information or object interaction information.

4.3 Effectiveness

In the last part, the questionnaire is focused on the qualitative on the effectiveness of the information. The question mentions about how influential in different information for the provision of health recommendation service. Likert scales is used from 1 (No, not influential at all) - 5 (Yes, it is critical). The result of this question is shown in Table 4. From the result of the first question, the respondents believe that the basic health information plays a vital role for diagnosis. Although the healthcare system can diagnose the disease with specific health information, the basic health information cannot find the cause of the disease.

Even tough the activity information obtained score in the second rank. It does not mean the activity information is not important. The SD value of the activity information is lower than the SD

Table 4. The qualitative of effectiveness of the information for provision of healthcare service

Information	Min	Max	Mean	SD
User's health	3	5	**4.33**	0.816
Human activity	4	5	4.16	**0.408**
Object interaction	3	4	3.50	0.547
Human location	2	4	3	0.632

of the basic health information. Thus, the importance of the activity information is closely to the basic health information because the activity information is used to analyze the cause of disease. It is useful for preventing or prognosis part in the healthcare system. The healthcare system can provide the health recommendation service for preventing or prognosis the disease to the home user based on the activity information analyzing.

The object interaction information and the human location information are in the third and last rank respectively. The respondents believe that these two information is hard to infer to the disease. However, if this information is integrated with other information such as activity information. It can improve the accuracy for selecting the appropriate service for the home user. For example, the "Computer" is turned on for a long time. The healthcare system cannot used this object interaction information to link with the "Tension-type headache" because the system cannot guarantee that home user is currently using the "Computer". Thus, the activity information can help to perceive this kind of knowledge.

5 Conclusion

In this paper, we introduced the CARE architecture that can provide the context-aware information in smart home. It is useful for analyzing the health recommendation service. The CARE architecture was designed as a human activity framework, and expressed the high accuracy of recognition when add two pieces of information: human posture information and AL^2. Based on the outputs of CARE architecture, the human behavior analysis system was built for summarization the human behavior information and the context-aware information in smart home. Four kinds of information were summarized and displayed to the home user and the physician. Moreover, the questionnaire was designed to evaluate the information, which was provided from the proposed system. The results of questionnaire showed that the respondents believe the activity information was relevant for diagnosis the disease. Although it was not the most important information, it was useful for the healthcare system. Since it can be used to find the cause of the problem and also can make the health recommendation service for prognosis.

Since the CARE architecture was designed based on the ontology concept, the advantage of ontology concept can be expressed in the future research. For example, the semantic web technology can be implemented for data sharing and processing in the area of healthcare.

References

[1] L C. De Silva, C. Morikawa, and I. M. Petra. State of the art of smart homes. *Eng. Appl. Artif. Intell.*, 25(7):1313–1321, 2012.

[2] S. Takahashi, S. Maeda, N. Tsuruta, and T. Morimoto. A home health care system for elderly people. In *Proceedings on KORUS 2003.*, pages 97–102 vol.2, 2003.

[3] K. Wongpatikaseree, A.O. Lim, Y. Tan, and H. Kanai. Range-based algorithm for posture classification and fall-down detection in smart homecare system. In *Proceedings on GCCE 2012*, pages 243–247, 2012.

[4] K. Wongpatikaseree, M. Ikeda, M. Buranarach, T. Supnithi, A.O. Lim, and Y. Tan. Location-based concept in activity log ontology for activity recognition in smart home domain. In *Semantic Technology*, LNCS, vol. 7774, pages 326–331. 2013.

[5] M. Buranarach, Y.M. Thein, and T. Supnithi. A community-driven approach to development of an ontology-based application management framework. In *Semantic Technology*, LNCS, vol. 7774, pages 306–312. 2013.

[6] Y. Tan. *Home Network Technologies for Smart Home.* Impress R&D, 2011. (in Japanese).

[7] I.M. Lopes, B.M. Silva, J.J.P.C. Rodrigues, J. Lloret, and M.L. Proenca. A mobile health monitoring solution for weight control. In *Proceeding of WCSP 2011*, pages 1–5, 2011.

[8] Y.C. Wu, P.F. Chen, Z.H. Hu, C.H. Chang, G.C. Lee, and W.C. Yu. A mobile health monitoring system using rfid ring-type pulse sensor. In *Proceeding of DASC 2009*, pages 317–322, 2009.

A Proposal of KOSA Model for Service Innovation

Michitaka Kosaka, **Hiroyuki Sakano, H.M. Belal, Qi Zhang and Kunio Shirahada**

Japan Advanced Institute of Science and Technology
kosa@jaist.ac.jp

Abstract

Innovations can improve the productivity and quality of services and lead to new services. Procedures to create services innovation have been a focus of discussion all over the world. So far, we proposed several methodologies such as KIKI model [1] or Knowledge space [2] for service innovation independently. However, these methodologies should be integrated effectively for service innovation. This research proposes KOSA (Knowledge Science Oriented Service Value Creation Architecture) model, which can express a standard procedure for creating service value. Several examples show the effectiveness of KOSA model.

Keywords: Services innovation, Knowledge science, KOSA model, Service value co-creation, KIKI model

1 Introduction

The importance of innovation and scientific investigation in the service sector has been discussed all over the world. Innovation in services should have an affect on productivity and quality and lead to new services. The importance of procedures for services innovation has been also discussed. So far, we have proposed several methodologies for service value creation independently, which are new service definitions [3], the service field [4], KIKI model [1], the knowledge space [2], the recursive approach [5]. The effectiveness of these methodologies was examined through several case studies [6]. However, these methodologies are not integrated for a standard procedure to create service innovation.

On the other hand, we already proposed the old KOSA (Knowledge Science Oriented Systems Approach) model for new business creation. However, this model did not consider customers' value creation. The KOSA model should be enhanced from the viewpoint of service value creation.

In this research, we try to integrate these methodologies into a new KOSA (Knowledge Science Oriented Service Value Creation Architecture) model as a standard procedure for service innovation. First, this paper reviews the previously proposed methodologies. Then, some examples are reported to show the effectiveness of these methodologies. Finally, these methodologies are integrated as the KOSA model.

2 Previous proposal for service innovation

We already made several proposals for service innovation. In this section, those are summarized briefly.

2.1 Two interpretation of service innovation

The term "service innovation" consists of "service" and "innovation" [3]. As the name implies, service innovation fuses the following two views.

(1) ***First view: Innovation in service industry***

The first view of service innovation means innovation in the service industry. This is a conventional view of service innovation where new technologies such as information technology or new hardware technology are employed to create new values or improve productivity in traditional service industries such as hotel and travel businesses.

(2) ***Second view: Innovation through service value creation***

The second view is innovation introduced through value co-creation between customers and providers from the viewpoint of SDL. Recent new services or new businesses fit this category, where new values in products or new services are created from the viewpoints of customer. This type of service innovation creates new business fields such as those created by Apple. It emphasizes new value creation by using information technology to help customers in the 21st century.

In this sense, the second view of service innovation is an important business issue with this

trend in the era of the Internet and globalization.

2.2 Definitions of service

The following three definitions of service are very useful in considering value of service in service innovation.

(1) ***A definition of service by a famous practitioner***

There are many Japanese traditional hot spring spas called "ryokan" where traditional Japanese hospitality (omotenashi) services are provided. Kagaya is one of the most well-known "ryokan". Sadahiko Oda, who is the chairman of Kagaya and is a well-known service practitioner, describes service as an activity that (a) provides professional techniques, (b) satisfies the customer, and (c) results in compensation [3]. These three factors in service are necessary conditions for successful service businesses and inspire fundamental philosophy in service science. According to his definition, most productive human activities are considered services.

(2) ***A definition of service by Kameoka***

He defined service as activities to support people or organizations for them to achieve their objectives [7]. He also defined service from the viewpoints of knowledge science, where service can be defined as a knowledge creation process of creating values for customers. A service business can be defined as providing created knowledge to customers. Therefore, service is related to knowledge creation.

(3) ***Service dominant logic***

Vargo and Lusch [8] proposed a new concept of service. This was called Service dominant logic (SDL). The most important point in SDL is that the determination of value in SDL differs from that in Goods dominant logic (GDL). The value of goods in GDL is determined by products, but the service value in SDL is determined by the customer on the basis of "value in use". Vargo and Lusch also stressed the importance of value co-creation with customers in SDL. The service value in SDL is determined by customers. Therefore, they are collaborators in creating service value and active participants in service creation. The major characteristic of SDL is that service value can be co-created by service providers and customers.

2.3 Previously proposed methodologies

The following 4 methodologies for service value creation were proposed for considering new services.

(1) ***Service field*[4]**

The value of services provided to customers generally depends on the efficiency of services to achieve customer objectives. Even though services are identical, their values differ in terms of how customers value the services, due to different customer characteristics in different situations (place, time, or cost). That is to say, service values depend on the contexts of the situation where a service is provided. We consider that a service value depends on the relationship between the service itself and its service field where the service is provided.

This service field concept is based on an analogy to the electromagnetic field. Even a good service will have no service value, if there are no service fields related to it. This is expressed as

(service value) = (service) x (service field)

where x means the relationship between the service itself and its service field. This indicates that service value depends on the service itself and its service field.

(2) ***KIKI model*[1]**

Service value depends on the relationship between the service and its situation, and a more suitable service can be provided if the service field is identified. From such considerations, We devised the following four steps for service value co-creation in B-to-B collaborations.

Step1(K1): Knowledge sharing in collaboration

Step2 (I1): Identification of the service field

Step3(K2): Knowledge creation for the new service idea

Step4(I2):Implementation of the new service idea

The above four steps in the service value co-creation process, which is KIKI model, can be described on a two-dimensional plane.

(3) ***Knowledge space*[2]**

The concept of knowledge space makes a bridge of collected knowledge, resources or experience between alliance partners. The combined competence of knowledge will then produce the required solution as well as co-create value for recipients. Thus, all knowledge of the firm and its partners, i.e. Alliance Company and recipients, enriches the knowledge space. By using this knowledge space, the firm becomes able to deliver a valuable solution and satisfy the customer.

(4) ***Recursive approach*[5]**

The recursive approach is used to assess customer experiences and it clearly identifies the requirements of recipients. Organizations can find appropriate services or products that satisfy customer expectations. These technical achievements were made possible by introducing a concept of Kalman filter that is useful for feedback control in system theory.

Organizations (service providers) have to analyze where and what the gaps are with the aim of meeting customer requirements or eliminating them between their expectations and those of the customers. Service providers can then enhance or innovate services or products to create further offerings based on analyses of the gaps. If this process is organized recursively or continuously by service providers then such gaps may reduce over time and finally be eliminated.

2.4 The old KOSA model for new business creation process

The old KOSA (Knowledge Science Oriented Systems Approach) model [9] was already proposed for new business creation innovation process based on the integration of system science and knowledge science, which was specified as follows. This model did not consider value creation for customers.

(1) ***Innovation process integrating system science and knowledge science***

The innovation process for new business development can be described as Figure 1. Usually, the innovation process for new business creation integrates system engineering factors (upper left) and knowledge science factors (lower right) for achieving the objective of new business development (an ideal state) from the initial stage.

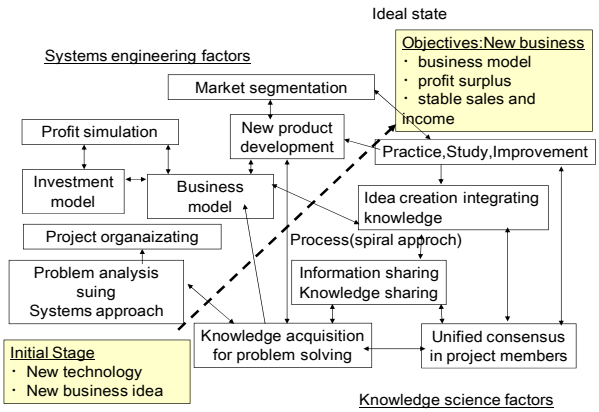


Figure 1 Innovation process

(2) ***Architecture of the old KOSA model***

By analyzing the structure of Figure 1, we could develop the old KOSA model as shown in Figure 2, which consists of 5 factors: (a) target vision, (b) knowledge space, (c) systems engineering, (d) knowledge science, (e) knowledge creation process management. In Figure 2, new knowledge for innovation is created in the knowledge spcae based on knowledge creation process management, which consists of SECI model, recursive approach, observability, and information potential (information field).

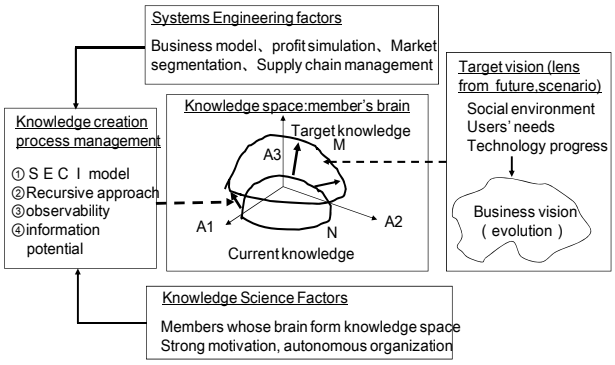


Figure 2 Architecture of KOSA model

(3) ***Failure information (estimation error) as innovation in Kalman filter***

Kalman filter is well known in the stochastic control theory. Kalman filter's concept is recursive approaching to the ideal state by reducing estimation error using recursive observation. In Kalman filter, innovation means estimation error and it is important information for expanding observation space. In the case of new business creation, failure information is a source of new knowledge creation in order to overcome failure at the current situation by expanding knowledge space. This characteristic can be explained by Figure 3.

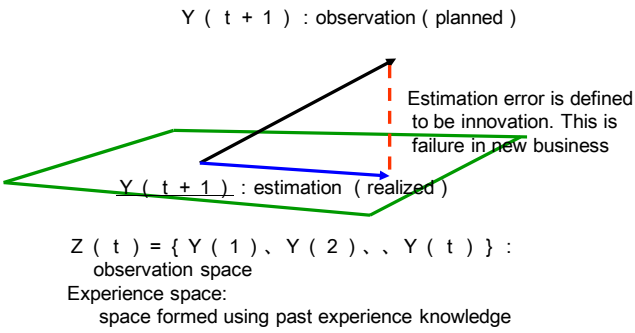


Figure 3 Estimation error as innovation

(4) ***Recursive approach***

Based on the consideration of (3), improvement of business or product specification should be done by creating new knowledge to reduce the failure. This process can be specified as Figure 4. This is feedback, spiral approach based on the experiences.

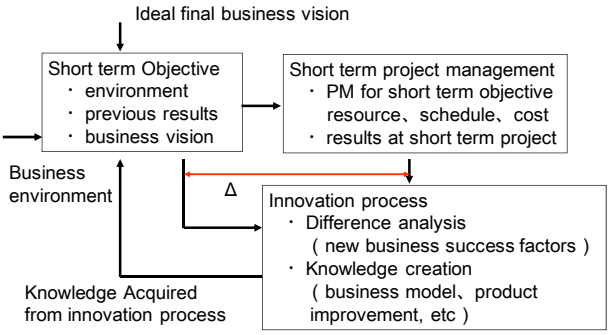


Figure 4 Feedback approach

(5) ***Effectiveness and weakness of the old KOSA model***

The effectiveness of the old KOSA model was evaluated through case studies such as a finger vein system business or an energy saving service business. However, the viewpoint of customer's value co-creation is not included in this KOSA model.

3 Effect of methodologies for service value creation in real business

We evaluated the effect of four methodologies for service innovation through analyzing four real business cases (Hand-held Terminal based Restaurant Ordering System, Electronic Dictionaries with Full Dictionary Contents, ThumBoard (Micro Keyboard) for PalmPilot (PDA), Electronic Dictionaries with E-book Reader); two successful ones and two unsuccessful ones respectively [6].

While the product planners then were not clearly aware of concepts of "value in use" or "value co-creation with customers", Table 1 shows what happened and what did not happen in this context among this company, customers, and their partner companies during the period of product planning and following product development processes.

We could draw some common elements and

Table 1 Case study of methodologies in real business

	Co-creation of Knowledge	Proposition of Value in use	Contribution by Outsiders	Primary Unit of Exchange	Business Results
Case 1 Restaurant Ordering System	Yes, many. But, most customers' know-how are masked	Yes, it is valid till today due to the customization for each customer	Yes, an advice from outsider in other industry was triggered	Service Value for accurate & quick order-taking at Restaurant Chain	Very Good. profitable. Created a market
Case 2 Full Contents E-Dictionary	Yes, a few. Some knowledge are protected as IP	Yes, it is valid till today though many of them are copied by others	Yes, an advice from engineer who did not know industry was triggered	Service Value for combination of full size dictionaries from other publishers	Very Good. Yet, soon Profit became lower
Case 3 ThumBoard® for PalmPilot®	Little though some IPs were registered by two companies	Little. Timing was too late due to slow down of PDA sales	None	Service Value for add-on function as small keyboard for PDA models	Very Poor. Timing was too late. Yet, was it really a reason?
Case 4 E-Dictionary with E-Book Reader	Almost none. NIH Syndrome between two companies	Small. Timing was too early (No Japanese E-Book Reader was ready)	None	Service Value f/ combination of dictionaries & e-book reader in one case	Poor. Timing was too early.

factors in successful and failed cases on Table 2. Among them, it appears that the cases in which knowledge co-creation process with customers (recipients of service) worked smoothly were mostly successful. Successful cases employed four methodologies for service innovation.

Table 2 Comparison of case studies

	Knowledge Space worked?	Service Field identified?	Knowledge Co-creation (e.g. KIKI) achieved?	Recursive Approach worked?
Case 1 Restaurant Ordering System	○	○	○	○
Case 2 Full Contents E-Dictionary	○	○	○	△
Case 3 ThumBoard® for PalmPilot®	×	×	×	×
Case 4 E-Dictionary with E-Book Reader	×	△	×	×

4 KOSA model for service innovation

4.1 Necessity of integrating methodologies into one framework for service innovation

From discussions in Section 3, the methodologies for service innovation are very effective for creating sevice innovation. These methodologies should be intergrated into one framework for creating service innovation systematically. On the other hand, the old KOSA model contains factors used in the methodologies, which are the knowledge space, recursive approach, SECI model and so on. Therefore, the old KOSA model should be enhanced to the new KOSA (Knowledge Science Oriented Service Value Creation Architecture) model by considering the viewpoint of service value creation. Also, these methodologies for service innovation should be integrated into this new KOSA model.

4.2 The new KOSA model for creating service innovation

The new KOSA model is shown in Figure 5. The service value creation process can be expressed by KIKI model. Methodologies for service

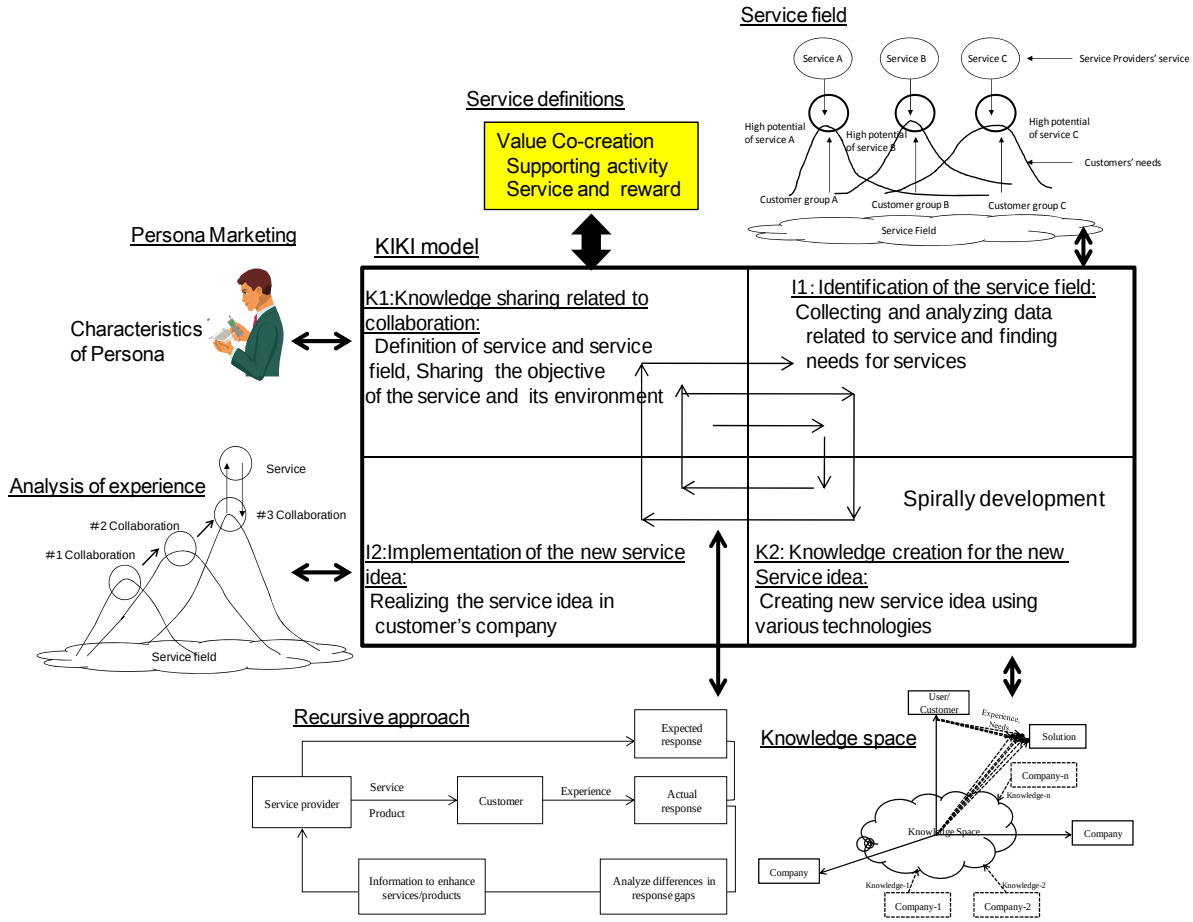


Figure 5 KOSA model architecture

innovation such as knowledge space or recursive approach are corresponding to each step K1, I1, K2, I2 in KIKI model. Persona marketing is employed in K1 to specify the customer characteristics. The service field is applied in I1 to identify the customer's need for value creation. The knowledge space is the framework for knowledge creation by integrating various knowledge and technologies in K2. The recursive approach is employed for enhancing service value in the spiral development. Also, an analysis of experience value on the provided service is applied to the spiral development. The new service development in KIKI model should consider the definitions of service and the characteristic of service "value in use".

5 Conclusion

Now, in the era of S-D Logic, working with customers and partners to co-create new knowledge for service value is essential. To make it happen, a new framework KOSA (Knowledge Science Oriented Service Value Creation Architecture) model is proposed, which can express a standard procedure for creating service value by integrating methodologies such as Knowledge space, KIKI Models, Service field and Recursive approach in Service System. The proposed KOSA model should be demonstrated to be effective though further researches.

Acknowlegement

This research was supported by Japanese Grants-in-Aid for Scientific Research (KAKENHI No.25240049).

References

[1] M. Kosaka, Q. Zhang, W. Dong and J. Wang, Service value co-creation model considering experience based on service field concept, *Proc. of IEEE International conference on service system and service management*, pp.724-729, Shanghai, China 2012.

[2] H. M. Belal, K. Shirahada and M. Kosaka, Knowledge space concept and its application for servitizing manufacturing industry, *Journal of Service Science and Management*, Vol.5, No.2, pp.187-195, 2012.

[3] M. Kosaka, A service value creation model and the role of ethnography, *An Ethnography of Global Landscapes and Corridors*, InTech, pp.109-130, 2012.

[4] M. Kosaka, K. Shirahada, and Y.Ito, A concept of service field in service systems for creating service value, *Proc. of the 4th Japan-China Joint Symposium on Information Systems* (JCIS2011), pp.35-40, 2011.

[5] H.M.Belal, K. Shirahada, and M. Kosaka, Value co-creation with customer through recursive approach based on Japanese Omotenashi service, *International Journal of Business Administration*, Vol.4, No.1, pp.28-38, 2013.

[6] H. Sakano and M. Kosaka, A consideration on success factors in product innovation from the viewpoints of value co-creation with customers, *IEEE International conference on service system and service management*, Hong Kong, 2013.

[7] A. Kameoka, *Service Science* (in Japanese), NTS, 2007.

[8] S.L. Vargo and R.F. Lusch, Evolving to a new dominant logic for marketing, *Journal of Marketing*, No.68, pp.1-17, 2004.

[9] M.Kosaka, A Methodology Integrating systems engineering and knowledge management for creating new business (in Japapnese), *Systems, Control and Information*, Vol.52, No.6, pp.29-35, 2008.

Value creation, co-production and co-creation in NGO-operated clothing industry

Md. Abul Kalam Siddike†* **Amna Javed†** **Youji Kohda†**

†School of Knowledge Science
Japan Advanced Institute of Science and Technology
Nomi, Japan
*Department of Information Science and Library Management
University of Dhaka, Bangladesh
{kalam,kohda,s1250207}@jaist.ac.jp

Abstract

The objective of this study is to conceptualize value creation, co-production and co-creation in non-government organization (NGO)-operated clothing industry. Secondly, we design value creation, co-production, and co-creation model for NGO-operated clothing industry. Thirdly, we demonstrate the model by taking the cases from Grameen UNIQLO and Aarong (two fast clothing companies in Bangladesh). Furthermore, we also conducted two interviews from Grameen UNIQLO. Finally, we give future directions for developing a unified model of value creation, co-production, and co-creation for clothing industry that will be utilized to other sectors.

In this paper we define value creation as a unidirectional process that flows from producers to customers to create value in the form of products. In traditional clothing industry, producers produce cloths for customers. We denote value co-production as a joint participation of customers and producers in the process of manufacturing and developing products and/or services. For NGO-operated clothing industry, customers design, process/arrange, and produce products/services by working together with producers and designers in clothing companies. Finally, we also define value co-creation as a bidirectional process that flows for a joint creation of value between customers and producers. In case of NGO-operated clothing industry, producers develop or create new products or services jointly with customers. Here, customers participate in the production process, advertising process, and selling process. That means customers play co-role as producers and customers.

Resutls from cases show that Grameen UNIQLO is engaged in community-oriented but Aarong engaged in market-oriented value creation, co-production, and co-creation.

Keywords: Value creation, Value co-production, Value co-creation, Service-dominant logic, Goods-dominant logic, NGO-operated clothing industry.

1 Introduction

Service-dominant Logic (SDL) is an alternative perspective to the traditional, Goods-dominant Logic (GDL) paradigm, which has been recognized as a potential theoretical foundation of service science [1]. GDL is the conventionally dominant logic of marketing and market which suggests that the purpose of economic activity is to make and distribute units of output (goods) which are embedded with utility (value) during production. On the other hand, SDL is a reoriented paradigmatic perspective on marketing and market which implies that service is the basis of all economic and social exchange; all businesses are service businesses and all economics are service economies. According to Vargo and Lusch [2, 3] service is the application of competences (knowledge and skills) by one entity for the benefit of another. In SDL, goods are a distribution mechanism for service provision.

In this study, first of all we review the literature for conceptualization of value creation, co-production, and co-creation in NGO-operated clothing industry. Secondly, we design value creation, co-production, and co-creation model for NGO-operated clothing companies. Thirdly, we demonstrate the model by taking the cases from Grameen UNIQLO and Aarong (two fast fashion companies in Bangladesh). Furthermore,

we also conducted two interviews from Grameen UNIQLO. Finally, we give future directions for designing a unified model for value creation, co-production, and co-creation in NGO-operated clothing industry, which will be utilized to other sectors.

2 Literature review

2.1 Value creation

Value creation is the central concept in service marketing. Different persons define value creation in various ways. Table 1 describes the definition of value creation.

Table 1. Definition of value creation

Authors	Definition of value creation
Prahalad and Ramaswamy [4]	Value creation is the process in which companies and customers play a distinct role of producer and consumer. It is a one way process that flows from firm to consumer, and controlled by firms only.
Ueda, Takenaka and Fujita [5]	It is an act of creativity in which the value for producer and consumer can be determined separately.
Gronroos and Voima [6]	It is an ongoing process that explains the customers' ability to extract value from products and other resources in use.
Hakanen and Jaakkola [7]	Benefit received from a combination of potential resources is value creation.

In this study, we define value creation as a unidirectional process that flows from producers to customers to create value in the form of products. In case of traditional clothing industry, producers produce cloths for customers.

2.2 Value co-production

In our review, the number of literatures on co-production was not so many. Table 2 denotes the definition of value co-production.

Table 2. Definition of value co-production

Authors	Definition of value co-production
Ordanini and Pasini [8]	Co-production is by using the knowledge of customers to maximize the service exchange benefits.
Jacob and Rettinger [9]	The interaction between buyer and seller for the purpose of creation of new product is co-production.

In our study, we denote value co-production as a joint participation of customers and producers in the process of manufacturing and developing products and/or services. In case of NGO-operated clothing industry, customers design, process/arrange, and produce products/services by working together with producers and designers in clothing companies.

2.3 Value co-creation

Value is co-created through the mutual effort of firms, employees, customers, stakeholders, government agencies, and other entities related to any given exchange. Therefore, the customer-manufacturer contact point plays an important role for value co-creation [10]. The SDL notion of value co-creation suggests that ''there is no value until an offering is used – experience and perception are essential to value determination'' [11]. Different researchers define value co-creation differently. Table 3 shows the definition of value co-creation.

Table 3. Definition of value co-creation

Authors	Definition of value co-creation
Ordanini and Pasini [8]	Value co-creation is the enhancement of value by using the resources and competencies of the business and customers.
Prahalad and Ramaswamy [12]	It is a joint creation of value and experiences by the company and customers having some necessary adjustments.
Ramaswamy [13]	Value generated from jointly developed products, services and experiences by companies and customers is value co-creation.
Prahalad and Ramaswamy [4]	It is a two way process that flows from firm to consumers and consumers to firm in return. Consumers have the right to impose their view of choice.
Ueda, Takenaka and Fujita [5]	It is an act of collective creativity that is experienced jointly by two or more people so the value for producer and consumer cannot be determined separately.
Vargo, Maglio and Akaka [14]	It is a continuous process of discovering unique resources and efficiencies by the participation

	of more than one service system.
Zhang and Chen [15]	The integration of customers for creation of value through co-creation activities and system capabilities.
Jacob and Rettinger [9]	Value created by the integration of firms offerings with other resources is co-creation.
Gronroos and Voima [6]	It is a process of mutual value expansion by the actions of both producer and consumer.

We define value co-creation as a bidirectional process that flows for a joint creation of value between customers and producers. In case of NGO-operated clothing enterprises, producers develop or create new products or services jointly with customers. Here, customers participate in the production process, advertising process, and selling process. That means customers play co-role as producers and customers.

In summary, while value creation occurs independently, value co-creation always accompanies value co-production, though the degree of "accompany" varies case by case.

3 Value creation, co-production, and co-creation model for NGO-operated clothing industry

3.1 Traditional clothing industry

Clothing is one of the very basic human needs. In general, clothing industry is now fully GDL-based. Traditional GDL-based clothing industry produces products for the customers and customers purchase products by the exchange of money. In this sense, value is measured by the products itself and the exchange of money. Figure 1 depicts the traditional GDL-based clothing industry for developed word.

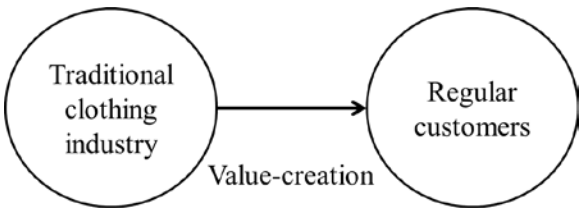


Figure 1. Traditional clothing industry for developed countries.

In the next section, we develop a model for NGO-operated clothing industry that connotes value creation, co-production and co-creation. Furthermore, we demonstrate the model by taking cases from two NGO-operated clothing companies and finally, we make comparison between the two cases.

3.2 NGO-operated clothing industry

Grameen UNIQLO [16] and **Aarong** [17] are two leading clothing companies in Bangladesh. These two companies look like belonging to GDL-based industry, but it produces and sells cloths with the aim of helping to address some of Bangladesh's most pressing social issues, including problems related to poverty, health sanitation and education. Figure 2 shows just an initial snapshot model for NGO-operated clothing industry in developing countries.

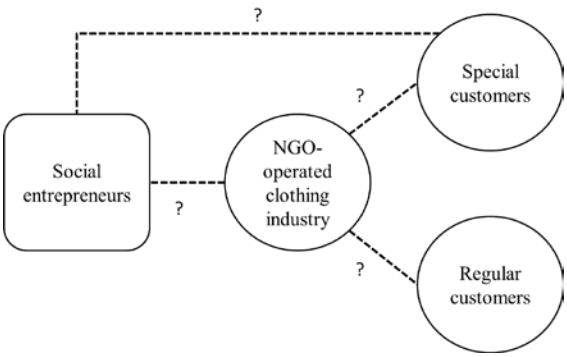


Figure 2. Initial snapshot model for NGO-operated clothing industry in developing countries

In figure 2, we can see that in NGO-operated clothing industry, there are both special and regular types of customers. The special customers who are the members of Grameen Bank and Brac (Bangladesh Rehabilitation Assistance Committee). Here, social enterpruners work with Grameen Bank and Brac. In both cases, the created values from NGO-operated clothing companies are being re-invested in social entrepreneurship business or social business for the well-being of special customers (the members of Grameen Bank and Brac).

3.3 Value creation, co-production, and co-creation model for Grameen UNIQLO

Firstly, Grameen UNIQLO is reducing poverty specially setting up jobs for Grameen ladies, who

are the special customers and members of Grameen Bank, employing them as sales assistants in the apparel retail industry. The Grameen ladies sell clothing door to door while explaining the key features of the products to customers within the network of eight million borrowers. Here, Grameen ladies are playing an important co-role as a producer as well as a customer and co-creating values. Secondly, it also reduces sanitation by producing and selling sanitary underwear and sweat pants. Finally, Grameen UNIQLO is raising literacy of children in Bangladesh through producing 'T-shirts with educational paints and electronic study books'. Figure 3 describes value creation, co-production and co-creation model for Grameen UNIQLO in Bangladesh.

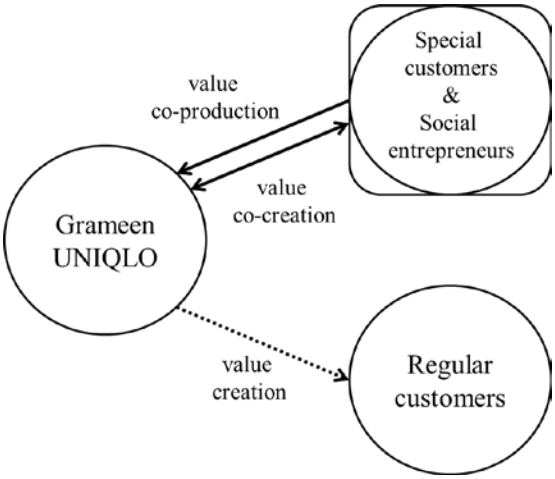


Figure 3. Value creation, co-production, and co-creation model of Grameen UNIQLO

In figure 3, we can see that producers (special customers) are jointly developing new products with customers (rural people) through the sharing of experiences, knowledge, and wisdom that indicate value co-production. Grameen UNIQLO mainly produces clothing for their eight million borrowers. It reinvests their profits for social business, educational program, health program, etc. that denote value co-creation. From figure 3, we also observe that producers (special customers) produce products for regular customers that indicate value creation.

3.4 Interview with Grameen UNIQLO's managers

We also conducted two interviews from Grameen UNIQLO. Both interviewees conformed that the borrowers of Grameen Bank work as producers as well as customers. One interviewee reported that rural women (special customers) sell cloths from door to door in rural areas or use their own homes as stores, where they explain product features and sell clothing in a warm and cordial setting. He also added that products are sold on consignment with paid commission based on sales. Another interviewee indicated that profits from the sales of the clothes will be reinvested in other social businesses. He also added that Bangladeshis are developing the business on their own, they have a direct hand in generating job opportunities, improving their lives and encouraging each other to move toward financial independence.

The interviewees reported that "The Value of $1" project is a new social business in which participation is as simple as purchasing an electronic study book for just one dollar. The one dollar spent to purchase an electronic study book of the Bengali language goes to a fund that subsidizes the production of T-shirts with educational prints and study books that will help raise the literacy of children in Bangladesh.

3.5 Value creation, co-production and co-creation model for Aarong

Aarong, is the local Mecca for ***deshi*** (local) handicraft. Aarong's product designs have brought consumer attention back to the products and styles that are indigenous to Bangladesh, its designers blending the traditional with the contemporary in a manner that has won instant consumer appeal, starting a revolution in trends that has now been taken up by countless other boutiques and stores. Sixty five thousands artisans and handicraft producers, as well as twenty five thousand independent cooperative groups and traditional family-based artisans who are the special customers, market their products and crafts through Aarong.

Potters, brass workers, jewellers, jute workers, basket weavers, handloom weavers, silk weavers, wood carvers, leather workers and various artisans with specialized skills from all over the country come to Aarong for marketing and support services. Figure 4 shows value creation, co-production and co-creation model for Aarong in Bangladesh.

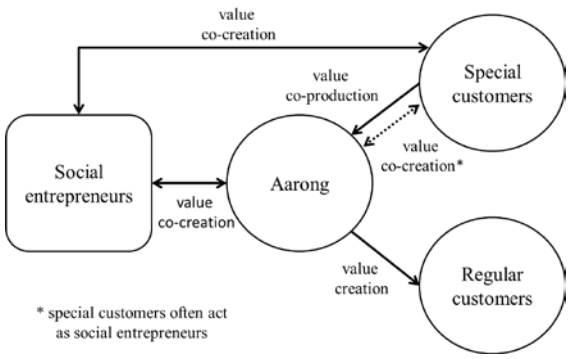


Figure 4. Value creation, co-production, and-co-creation model of Aarong

In figure 4, we can see that the special customers (rural people who are the members of Brac) along with designers from Aarong are producing products mainly for the regular customers that indicate co-production. Here, special customers who are the borrowers of Brac and the designers from Aarong are taking part in the production process by sharing their knowledge, experiences and wisdom. But, Aarong is mainly producing products for regular customers that denote creation of value. On the other hand, Aarong also re-invests the profits for social entrepreneurship and special customers for their social business that indicate the co-creation value.

3.6 Comparison of Grameen UNIQLO and Aarong cases

These two clothing enterprises of Bangladesh are not only producing and selling cloths, but also they are creating job opportunities for rural women, reducing sanitation, raising literacy, providing health services. The co-created values from these two companies are re-invested to social business, health care, educational programs and empowerment of rural women in Bangladesh.

Through the comparison of two cases, Grameen UNIQLO mainly co-creates values jointly with customers. That means producers and customers are interacting together and producing and developing new products through the sharing of their knowledge, experiences, and wisdom. But in case of Aarong, they also co-create values very nominally. Because, Aarong‘s main customers are regular customers. That means Aarong mainly produces products for selling to regular customers and they re-invest their profits for health services, educational programs, as well as poverty eradication and empowerment of rural women.

In short, Grameen UNIQLO is engaged in community-oriented value co-creation and Aarong in market-oriented value co-creation.

4 Conclusion

Values are creating through the joint interaction of producers and customers by sharing their knowledge, experiences, and wisdom. In case of Grameen UNIQLO, special customers are taking part in the production and development of products. But in case of Aarong, producers are the special customers who are producing cloths for regular customers. Finally, the created values from these two companies also re-invest to social enterprunersship business which includes educational program, health services, and empowerment of rural women.

Value creation, co-production, and-co-creation model in NGO-operated clothing industry is unique in nature. In future, we will conduct qualitative research to justify the proposed model. Finally, the new model in NGO-operated clothing industry would be unique in nature that will broaden the dimension of SDL concept, and it will be applicable to other sectors.

Acknowledgement

The study is funded by Ministry of Education, Culture, Sports, Science and Technology (Monbukagakusho) scholarship, at Japan Advanced Institute of Science and Technology (JAIST), Japan. Therefore, the Monbukagakusho is greatly acknowledged for financial support.

References

[1] Vargo, S. L., Lusch, R. F. and Akaka, M. A. Advancing service science with Service-Dominant Logic: clarifications and conceptual development. In ‘Handbook of Service Science’, (eds.) Maglio, P. P., Kieliszewski, C. A. and Spohrer, J. C., 133-158. Springer, 2010.

[2] Vargo, S. L. and Lusch, R. F. Evolving to a new dominant logic for marketing, *Journal of Marketing*, 68(January): 1-17. 2004.

[3] Vargo, S. L. and Lusch, R. F. Service-dominant logic: what It Is, What It Is Not, What It Might Be. In The Service-Dominant Logic of Marketing: Dialog, Debate and Directions, (eds) R. F. Lusch and S. L. Vargo, 43-56. M. E. Sharpe Inc., Armonk, 2006.

[4] Prahalad, C. K. and Ramaswamy, V. Co-creating unique value with customers. *Strategy & Leadership*, 32(3): 4-9, 2004.

[5] Ueda, K., Takenaka, T. and Fujita, K. Toward value co-creation in manufacturing and servicing. *CIRP Journal of Manufacturing Science and Technology*, 1(1): 53-58, 2008.

[6] Gronroos, C. and Voima, P. Critical service logic: making sense of value creation and co-creation. *Journal of the Academy of Marketing Science*, 41(2): 133-150, 2013.

[7] Hakanen, T. and Jaakkola, E. An empirical investigation into value co-creation in b-to-b service networks. 2013.

[8] Ordanini, A. and Pasini, P. Service co-production and value co-creation: the case for a service-oriented architecture (SOA). *European Management Journal*, 26(5): 289-297, 2008.

[9] Jacob, F. and Rettinger, B. The role of customer co-production in value creation. 2010.

[10] Belal, H., Shirahada, K. and Kosaka, M. Knowledge Space Concept and Its Application for Servitizing Manufacturing Industry. *Journal of Service Science and Management*, 5(2): 187-195, 2012.

[11] Vargo, S. L. and Lusch, R. F. Service-dominant logic: What It Is, What It Is Not, What It Might Be. In The Service-Dominant Logic of Marketing: Dialog, Debate and Directions, (eds) R. F. Lusch and S. L. Vargo, pp. 43–56. M.E. Sharpe Inc., Armonk, 2006.

[12] Prahalad, C. K. and Ramaswamy, V. Co-creation experiences: the next practice in value creation. *Journal of Interactive Marketing*, 18(3): 5-14, 2004.

[13] Ramaswamy, V. Co-creation of value-towards an expanded paradigm of value creation. *Marketing Review ST. Gallen*, 26(6): 11-17, 2009.

[14] Vargo, S. L., Maglio, P. P. and Akaka, M. A. On value and value co-creation: a service systems and service logic perspective. *European Management Journal*, 26(3): 145-152, 2008.

[15] Zhang, X. and Chen, R. Examining the mechanism of the value co-creation with customers. *International Journal of Production Economics*, 116(2): 242-250, 2008.

[16] Grameen UNIQLO. About Grameen UNIQLO. 2013. Available at: http://www.grameenuniqlo.com/en/about/.

[17] Aarong. About Aarong. 2013. Available at: http://www.aarong.com/aboutus.php.

Capacity Maximization via Selection of Uniformly Distributed User for Uplink MIMO Communication over η-μ Fading Channel

Theerasak Sangyam† **Chalie Charoenlarpnopparut**† **Prapun Suksompong**†
Kamol Kaemarungsi‡ **Kazuhiko Fukawa***
†School of Information, Computer, and Communication Technology, SIIT, TU, Thailand
‡Embedded System Laboratory, NECTEC, NSTDA, Thailand
*Signal Processing Laboratory MCRG, Tokyo Tech, Japan
th.sangyam@gmail.com,
{chalie, prapun}@siit.tu.ac.th,
kamol.kaemarungsi@nectec.or.th,
fukawa@radio.ss.titech.ac.jp

Abstract

In this work, our objective is to analyze the performance, by deriving the average channel capacity, of the uplink MIMO communication system when η-μ fading channel is applied comparing to Rayleigh fading channel. In addition, the uniformly distributed user locations and user selection with capacity maximization are also used in the simulation. We are also investigating on 4 main schemes which are single-user (SU), multi-user (MU), single-input single-output (SISO) and multiple-input multiple-output (MIMO). The results shows that the performance of uplink communication system over η-μ fading channel is better than Rayleigh fading channel in SU scheme but worse in MU scheme. However, the increasing in the number of antennas and the number of user yield the better performance.

Keywords: Capacity maximization, Uplink communication, Uniformly distributed user, η-μ distribution

1 Introduction

Social network, e-commerce and video streaming have be-come a part of our daily life, which bring us to an emerging technology of wireless communication. While the number of users is increasing rapidly every day, the amount of data to be transmitted is also increased, especially for video streaming or video conference. Wireless communication has been used to substitute wired communication since it has lower infrastructure cost and takes less space. Moreover it is very convenient due to its mobility. The requirement of wireless communication system design is to achieve the highest performance i.e., reliability, capacity and range. For single-input single-output (SISO) system, there is strong tradeoff between reliability, capacity and range. On the other hand, multiple-input multiple-output (MIMO) system can improve these three parameters simultaneously. In addition, MIMO system increases the spectral efficiency and reduces the effects of the fading due to the increased diversity [1].

Single-user (SU) MIMO scheme has been studied in many wireless communication standards as well as IEEE 802.11n standard. In the past decade, multi-user (MU) has been considered as an interesting transmission technique for MIMO wireless communication. Numerical results from [2] show that MU MIMO scheme yields lower collision probability and shorter delay performance than SU MIMO scheme. Furthermore, in other cases of high SNR values and a small number of contending stations, MU MIMO also yields better throughput performance. In IEEE 802.11ac standard, only downlink MU MIMO precoding scheme has been implemented. Uplink MU MIMO precoding scheme is still not included. However, its application is necessary as shown in the usage model of IEEE 802.11ac standard from 802.11ac Taskgroup (TGac) which includes wireless display, distribution of HDTV, rapid upload/download, backhaul, outdoor campus/auditorium and manufacturing floor [3].

Most of the research topics on wireless communication, especially for wireless local area network (WLAN) are mainly focused on downlink transmission scheme as in [4–8]. This re-

search poses a challenge to authors when trying to find any related works and practice. Currently, there are only a few number of collective works dealing with uplink transmission scheme. As it is necessary to implement the uplink MU MIMO precoding scheme, user selection for uplink MU MIMO system and uniformly distribution of users in a square room is studied in this work. By comparing to the previous work [9] which investigated only on Rayleigh fading channel, we extend our work to η-μ fading channel.

The remaining of this paper is organized as follows. System model and fading model are provided in section 2 and 3 respectively. Simulation result shows in section 4. Finally, we summarizes the paper into section 5.

2 System Model

An efficient uplink multiuser MIMO protocol in IEEE 802.11 WLANs was studied in [10] and two main problems were considered. They are synchronized transmission among the stations and spatial compatibility between the transmitting stations. We choose to base our solution on the scheme discussed above, whose one of the main objectives is to increase the network throughput.

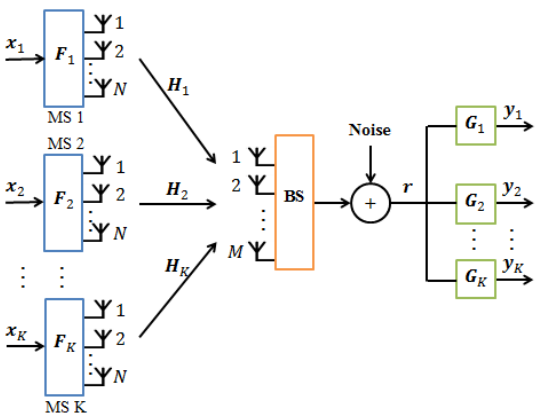


Figure 1. Uplink MU MIMO system.

Figure 1 shows the uplink MU MIMO system from [11] with a based station (BS) with M receive antennas and K mobile stations (MS) or users each with N transmit antennas. The users are allowed to transmit data simultaneously. $\boldsymbol{x}_i$ is the input data stream vector of length N from user i; $\boldsymbol{F}_i$ is $N \times N$ linear precoder; $\boldsymbol{H}_i$ is $M \times N$ uplink MU MIMO channel matrix between user i and BS; $\boldsymbol{n}$ is zero mean white complex Guassian noise vector. Then, the received signal vector $\boldsymbol{r}$ of length M at the BS can be written as

$$\boldsymbol{r} = \sum_{i=1}^{K} \boldsymbol{H}_i \boldsymbol{F}_i \boldsymbol{x}_i + \boldsymbol{n} \tag{1}$$

In [10], the BS calculates the channel state information (CSI) upon the receiving pilot signals which are sent from all users that request to transmit data and then selects the group of users that give the maximum channel capacity from all possible groupings. The capacity is calculated by

$$C_g = \log_2 \det \left(\boldsymbol{I}_{KN} + \boldsymbol{H}_g \boldsymbol{H}_g^H \right) \tag{2}$$

where $(\cdot)^H$ denotes for Hermitian transposition; $\boldsymbol{I}_{KN}$ is $KN \times KN$ identity matrix; and $\boldsymbol{H}_g$ is $KN \times M$ channel matrix of user group g which can be found by

$$\boldsymbol{H}_g = \left[\frac{P_1}{N\sigma_1^2} \boldsymbol{h}_1^T \ldots \frac{P_i}{N\sigma_i^2} \boldsymbol{h}_i^T \ldots \frac{P_K}{N\sigma_K^2} \boldsymbol{h}_K^T \right]^T \tag{3}$$

where $\boldsymbol{h}_i^T$ is $M \times N$ channel coefficients between user i and BS; σ_i^2 is the noise power and P_i is the transmission power; and $(\cdot)^T$ denotes the transpose operation.

Similar to [10], a group of users which gives the highest channel capacity among all possible $C(R, K)$ groups is allowed to transmit the data. Here, R is the number of all users requested to transmit data. In our simulation, a room size of 5×5 square meters is used while all MSs are located on the same level, 1 meter below the ceiling while a BS is installed at the center of the room. The rest of this section explains how to calculate capacity.

Reference SNR is assumed in the range from 0 dB to 20 dB with the incremental step of 1 dB. This reference SNR is located at 1 meter below the BS i.e., it is on the same level as the MSs. From this value, reference power can be obtained when noise power σ_i^2 is assumed to be 1. Then random locations of each user are uniformly generated in the square room. So the actual transmission power of each user can be calculated using simplified path loss model (4) from [12]:

$$P = P_{ref} \beta \left[\frac{d_0}{d} \right]^{\gamma} \tag{4}$$

where P and P_{ref} are transmission and reference power respectively. γ is the path loss exponent and β is the unitless constant. d_0 and d are a reference distance for the antenna farfield and transmission distances accordingly.

Channel $\boldsymbol{h}_i^T$ is also generated randomly depending on the fading model for each user and its size is depending on the number of transmit and receive antennas. In this paper, the number of transmit antennas are the same for every user. With assumption of noise power is equal to 1, the channel matrix $\boldsymbol{H}_g$ can be calculated from (3). Finally, the channel capacity C_g can be calculated by substitute $\boldsymbol{H}_g$ into (2). Then the channel capacity is calculated repeatedly 100,000 times. Then those values are averaged for each reference SNRs.

3 Fading Model

3.1 Rayleigh Distribution

This fading is most commonly and widely used as it is the statistical model for the effect of a propagation environment for non-line-of-sight (NLOS) wireless communication. The probability density function (pdf) of Rayleigh distribution , $f_R(r)$, is

$$f_R(r) = \frac{2r}{\sigma^2} e^{-r^2/\sigma^2}, r \geq 0 \tag{5}$$

where σ^2 is the variance. The Rayleigh distribution has an independent Uniformly distributed phase.

3.2 η-μ Distribution

The η-μ distribution is a general fading distribution for NLOS applications. It is more flexible because it uses two parameters to adjust behavior of the distribution.

It appears in 2 different formats. Format 1, the in-phase and quadrature components of the fading signal within each cluster are assumed to be independent from each other and to have different powers. $0 < \eta < \infty$ is the scattered-wave power ratio between the in-phase and quadrature components of each cluster of multipath. Format 2, the in-phase and quadrature components of the fading signal within each cluster are assumed to have identical powers and to be correlated with each other. $-1 < \eta < 1$ is the correlation coefficient between the scattered-wave in-phase and quadrature components of each cluster of multipath. In this work and the rest of this paper, we consider only on Format 1.

For a fading signal with envelope R and normalized envelope $P = R/\hat{r}, \hat{r} = \sqrt{E(R^2)}$ being the rms value of R, the η-μ pdf, $f_P(\rho)$, can be found in [13] as

$$f_P(\rho) = \frac{4\sqrt{\pi}\mu^{\mu+\frac{1}{2}}h^\mu}{\Gamma(\mu)H^{\mu-\frac{1}{2}}} \rho^{2\mu} e^{-2\mu h \rho^2} I_{\mu-\frac{1}{2}}(2\mu H \rho^2) \tag{6}$$

and the phase pdf, $f_\Theta(\theta)$, can be found in [14] as

$$f_\Theta(\theta) = \frac{(h^2 - H^2)^\mu \Gamma(2\mu) |sin(2\theta)|^{2\mu-1}}{2^{2\mu}\Gamma^2(\mu)(h + Hcos(2\theta))^{2\mu}} \tag{7}$$

where $h = \frac{2+\eta^{-1}+\eta}{4}$ and $H = \frac{\eta^{-1}-\eta}{4}$ in Format 1; $\Gamma(\cdot)$ is the Gamma function; $I_v(\cdot)$ is the modified Bessel function of the first kind and order v; and $\mu > 0$ is given by

$$\mu = \frac{E^2(R^2)}{2V(R^2)} \left[1 + \left(\frac{H}{h}\right)^2\right] \tag{8}$$

4 Simulation Result

Comparison of the performance over Rayleigh and η-μ fading channel with $\eta = 1$ and $\mu = 0.5$ which is equivalent to Rayleigh fading distribution and $\eta = 2$ and $\mu = 1$ which used in [15] are show as Figures 2 to 5.

Figures 2 and 3 show the simulation results of uplink SU (1 user) communication scheme with better result when $\eta = 2$ and $\mu = 1$ while Figures 4 and 5 are the simulation results of uplink MU (4 users) communication scheme with worse result when $\eta = 2$ and $\mu = 1$ comparing to Rayleigh fading.

On the other hand, Figures 3 and 5 show the simulation results of uplink MIMO (2 transmit and 2 receive antennas) communication scheme which yield better results than the simulation results of uplink SISO (1 transmit and 1 receive antenna) communication scheme as shown in Figures 2 and 4.

5 Conclusion

From the previous section, we may conclude that the performance of uplink communication over η-μ fading channel is achieved in SU scheme

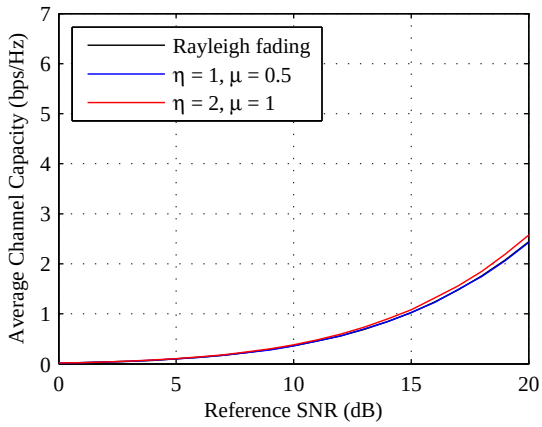


Figure 2. Average Channel Capacity of SU-SISO Scheme

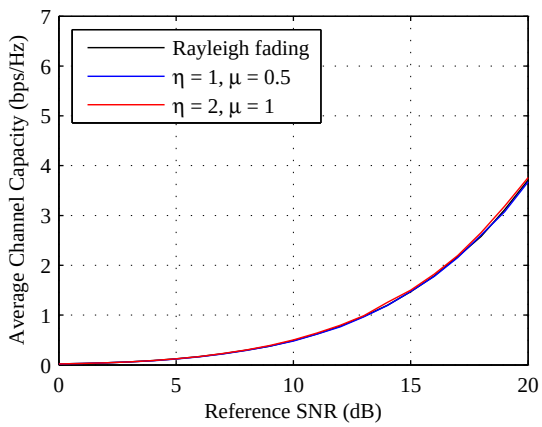


Figure 3. Average Channel Capacity of SU-MIMO Scheme

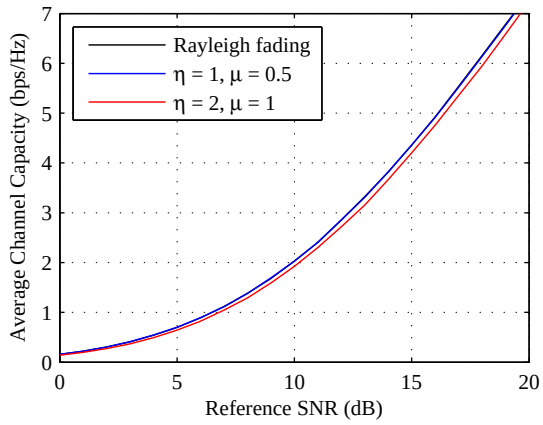


Figure 4. Average Channel Capacity of MU-SISO Scheme

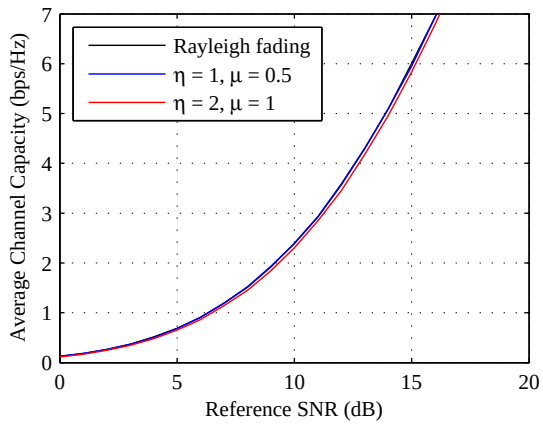


Figure 5. Average Channel Capacity of MU-MIMO Scheme

while MU scheme yields a little worse results comparing to Rayleigh fading channel. Nevertheless, the increasing in the number of antennas and the increasing in the number of user give the better capacity as shown in [9].

6 Future Work

In this work, we only used some specific value of η and μ. Thus, in the future work, we may vary these values to analyze the behavior or to find the best appropriate value according to our model.

Acknowledgment

This work was supported by the NTC Telecommunication Research lab, the National Research University Project of Thailand, Office of the Higher Education Commission, TRIDI Advanced Research Grant, NSTDA, SIIT and TAIST-Tokyo Tech.

References

[1] L. Hanzo, Y. Akhtman, L. Wang, and M. Jiang. *MIMO-OFDM for LTE, WiFi and WiMAX(Coherent versus Non-coherent and Cooperative Turbo Transceivers)*. IEEE Press and John Wiley & Sons, Ltd., 2010.

[2] Hu Jin, Bang Chul Jung, Ho Young Hwang, and Dan Keun Sung. Performance comparison of uplink WLANs with single-user and multi-user MIMO schemes. In *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, pages 1854 –1859, 31 2008-Apr. 3 2008.

[3] Rolf de Vegt. 11-09-0161-02-00ac-802-11ac-usage-model-document. Technical report, TGac, 2009.

[4] Ruizhi Liao, B. Bellalta, C. Cano, and M. Oliver. DCF/DSDMA: Enhanced DCF with SDMA downlink transmissions for WLANs. In *Internet Communications (BCFIC Riga), 2011 Baltic Congress on Future*, pages 96 –102, Feb. 2011.

[5] T. Hiraguri, T. Ogawa, T. Ueno, K. Nishimori, and K. Jin'no. Transmission queuing schemes for improving downlink throughput on IEEE 802.11 WLANs. In *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*, pages 1 –5, 31 2011-Aug. 4 2011.

[6] Jianbo Wu, Zhiyi Yin, Jinbao Zhang, and Wei Heng. Physical layer abstraction algorithms research for 802.11n and LTE downlink. In *Signals Systems and Electronics (ISSSE), 2010 International Symposium on*, volume 1, pages 1 –4, Sep. 2010.

[7] W. Hardjawana, B. Vucetic, and Yonghui Li. Cooperative precoding and beamforming for co-existing multi-user MIMO systems. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1 –5, Jun. 2009.

[8] T. Tandai, H. Mori, and M. Takagi. Cross-layer-optimized user grouping strategy in downlink multiuser MIMO systems. In *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pages 1 –6, Apr. 2009.

[9] S. Theerasak, C. Chalie, S. Prapun, Kaemarungsi Kamol, and Fukawa Kazuhiko. Performance of uplink multi-user mimo wifi with capacity maximizing user selection and uniformly distributed user locations. In *The third international conference on information and communication technology for embedded systems 2012, ICICTES 2012*, Jan. 2012.

[10] T. Tandai, H. Mori, K. Toshimitsu, and T. Kobayashi. An efficient uplink multiuser MIMO protocol in IEEE 802.11 WLANs. In *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, pages 1153 –1157, Sep. 2009.

[11] Chia-Chang Hu and Chung-Lun Yang. Combined transceiver optimization for uplink multiuser MIMO with limited CSI. In *International Scholarly Research Network, ISRN Signal Processing*, volume 2011, 2011.

[12] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[13] M.D. Yacoub. The κ-μ distribution and the η-μ distribution. *Antennas and Propagation Magazine, IEEE*, 49(1):68–81, 2007.

[14] D.B. da Costa and M.D. Yacoub. The η-μ joint phase-envelope distribution. *Antennas and Wireless Propagation Letters, IEEE*, 6:195–198, 2007.

[15] Apiwat Aksornsart and Prapun Suksompong. Performance gain of antenna selection for mrc over η-μ simo fading channel. In *The fourth international conference on information and communication technology for embedded systems 2013, ICICTES 2013*, Jan. 2013.

An Analysis of Major Influencial Factors for Innovation in Thai Automotive Industry

Narongsak Pongsathornwiwat† Chawalit Jeenanunta‡* Thanaruk Theeramunkong * Hyunh Van-Nam**

† , ‡, * Sirindhorn International Institute of Technology, Thammasart University, Thailand
**Japan Advanced Institute of Science and Technology
{npongsathornwiwat}@yahoo.com

Abstract

Innovation is one of the main sources of competitive advantage and essential for company growth. Companies put great effort in creating improving and introducing innovations. This paper presents a study on innovation in Thai automotive industry and determines the influential factors of innovation. The survey was conducted to explore influential factors, including supplier driven, technological driven, customer driven, organizational capabilities, competitors' pressure, organizational learning and learning of study of competitors which have impacts on innovation in the automotive industries. The results were analyzed by using a multiple linear regression technique and factor analysis. We find that the keys positively effecting to innovation include supplier driven, technological driven, customer driven, organizational capabilities and organizational learning; while, on the other hand, competitors' pressure and learning of study by competitors are negatively affecting to innovation in Thai automotive industries. And, technological driven is a key factor orienting innovation in Thai automotive industries.

Keywords: Innovativeness, Innovative capabilities, Innovation performance, Thai automotive firms, multiple linear regressions

1 Introduction

Due to globalization and sustained demand growth, the automotive industries are focusing more to sustain their customer base and to enhance profit opportunities; so, they must manage successive competitive advantage such as innovations effectively [3, 4, 6, 7].

Innovation is becoming the key factor enhancing the firm businesses to stay competitiveness by introducing and improving the process capabilities with something new such as a novel idea, method or device [1, 2, 12, 13]. It can be concluded that innovations play an important role in automotive industry. As social environments, economics and businesses change rapidly, many organization or firms has to create or invent new product, services and process in response to its customers' needs and pressure [14, 15]. In addition, due to declining of limited natural resources, any business that has to rely on natural resources must improve their business for their survival and sustainability. Thus, innovation is one of the major factors in a successful business [6, 7].

There are many studies investigate the factor influencing innovations in various industries such as semiconductor and electronics industries and also, high-tech industries [20]. However, there were few studies which explored factors influencing firm's innovation especially in Thai automotive industry. It is crucial, practical and interesting to investigate on how automotive firms can innovate and stay competitive. These reasons lead our attention to examine the appropriate factors influencing innovation in automotive industries established in Bangkok, Thailand so as to hopefully provide a understanding on their innovational situation. Therefore, the main objective of this study are qualitatively to examine what are the factors/attributes indicating the innovation in Thai automotive industry and how they affecting the innovation.

2 Literature review and related work

2.1 Innovation theory

Innovations assist firms and businesses in the way of increasing its performance as they indicate the positive and negative effects of the business model. The general definition of an innovation is the implementation of a new or significantly improved product/service, or process, a new marketing method, or a new organizational method in business practices [15]. Besides, the decisions of innovation determinants are the important processes of additional products or services on offer to the desired customer effectively. Therefore, the factors influencing innovation in manufacturing industries are needed to make more clearly understanding [3, 7]. According to Oslo Manual (OECD 2005), it can be categorized innovation into 4 types consisting of product innovation, process innovation, marketing innovation, and organizational innovation. The product/service innovation is the introduction of a new product including essential improvements, technical specification, components, materials, and other functional characteristics. For process innovation, it is the implementation of a new production method or delivery method (mainly focus on processes). Then, marketing innovation is the new way of marketing the product, also includes new market and mainly focus on latest marketing processes. And, the last one is organizational innovation. It denotes new organizational method in the firm's business practices, workplace organization or external relations.

Moreover, the level of innovation can be categorized into 2 types which are radical innovation and incremental innovation. Radical innovation brings about a non-routine activities change or major change; whereas incremental innovation is routine activities change or minor change [7, 15, 19].

2.2 Influential factors for innovation

There are many characteristics of the firms that play in orienting innovation within the manufacturing industries, especially in automotive industry [7, 9, 10]. The automotive industry sector is characterized by a number of key and unique characteristics from the perspective of product features and the sector's structure, where innovative practices are developed in response to the economic pressures and customer requirement, driving the evolution of the chain and encourage greater horizontal and vertical coordination [3, 4]. With the objective of manufacturing and supplying a specific product or component, or locating new enterprises, all enterprises in the manufacturing cooperate as synergetic unit to pursue for success [6]. Besides, a key successful factor driving towards innovation is customers. The linkage consists of linkage with stakeholders, skills of top management, human resource capabilities. However, not all the firms are successful in the innovative improvement [19]. The successful innovative implementation needs support and the willingness from all stakeholders, especially the support from management level and the quality of employees. Moreover, the technological capability is another important factor that contributes to creating the competitive advantage of the firm [7].

There are many factors to determine the impact on innovation used in this study. All the factors were reviewed from recent published articles [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16]. Hence, the present study will examine whether these factors positively influence the Thai automotive businesses' innovation.

3 Research Methodology

As the related literature review and the existing theory, the design of the scale and variables used in the present study are conducted for measuring these variables in quantitative five-point likert scale ranging from strongly agreed to not familiar by using the structured questionnaires that were sent in person, mails, and online for surveying the data. The process of this study follows the guideline in the revised OECD 'Oslo Manual'[15] focusing on innovation and automotive industry in developing country in case of Thailand. The questionnaire is divided into four sections; that is, firm's characteristics, factors promoting firm's innovation capabilities, measurement of innovation performances, and the measurement of the firm's performances from innovation developed. In this study, we show the analytical section especially the part of factors promoting firm's innovation.

This study had set up the assumption in believing that there are positive relationships between innovation factors and innovation measurement. After analyzing with factor analysis technique, there are seven main factors considered as independent variables for this study,

which are supplier driven, technological driven, customer driven, organizational capabilities, competitors' pressure, learning by studied from other companies and organizational learning. The description of each construct, which is supplier-driven measuring the ability of firms to work closely with appropriated supplier to initiate innovation. Technological driven measures how firms acquire the new technology both hardware and software to develop innovation. Customer driven measures the ability of firms to know and respond to satisfy the customer. Organizational capabilities measures how firms can have potential of abilities and skills to build innovation. Customers' pressure investigates the relationship and interaction with external parties for innovation sharing. Learning by studied from other companies and organizational learning measure how firms can learn and adopt knowledge to innovate competitive advantage. Innovation measurement represented as dependent variables that were affected by those seven factors. The next step is to measure the impacts or outcomes from innovation to the firm's performance with the assumption to believe that there is a positive relationship between innovations and the firm's performance. Therefore, the assumption leads to the following hypotheses;

Hypothesis 1: Supplier driven has a positive relationship on innovation development.

Hypothesis 2: Technological driven has a positive relationship on innovation development.

Hypothesis 3: Customer driven has a positive relationship on innovation development.
Hypothesis 4: Organizational capabilities have a positive relationship on innovation development.

Hypothesis 5: Competitors' pressure has a positive relationship on innovation development.

Hypothesis 6: Learning by studied from other companies has a positive relationship on innovation development.

Hypothesis 7: Organizational learning has a positive relationship on innovation development.

The concept of this model is shown in Figure 1. This research looks at automotive industries, and these factors that influence the innovation.

4 Data collection and analysis

4.1 Data collection

A total of 2,050 firms of automotive businesses had received by postal mails, online and in person with 174 returned represent 8.49 % in the respond rate. According to the personal information, it is shown that 29.31 % of the samples are multi-national companies, 24.14 % are joint venture companies, and 46.55 % are the most domestic companies. Besides, the most types of business are tier 2 and 3 by 83.9 %, followed by tier 1 and assembler with 8.05 % equally.

It is shown that the innovation factors came from multinational companies rather than joint ventures and domestic companies. As for the size classification of automotive companies; 46.58 % of the samples are large-sized companies with fixed assets available for more than 200 million baht, followed by 37.36 % for medium- sized companies that have 50 to 200 million baht; small-sized companies with less than 50 million baht amounting to 16.09 %. The number of employees classification of companies; automotive companies that have less than 50 employees are amounted to 10.92 %, companies that have more than 50 but less than 200 employees are amounted to 25.86 5%, more than 200 but less than 500 are calculated as 35.63 %, more than 500 employees account for 13.22 %.

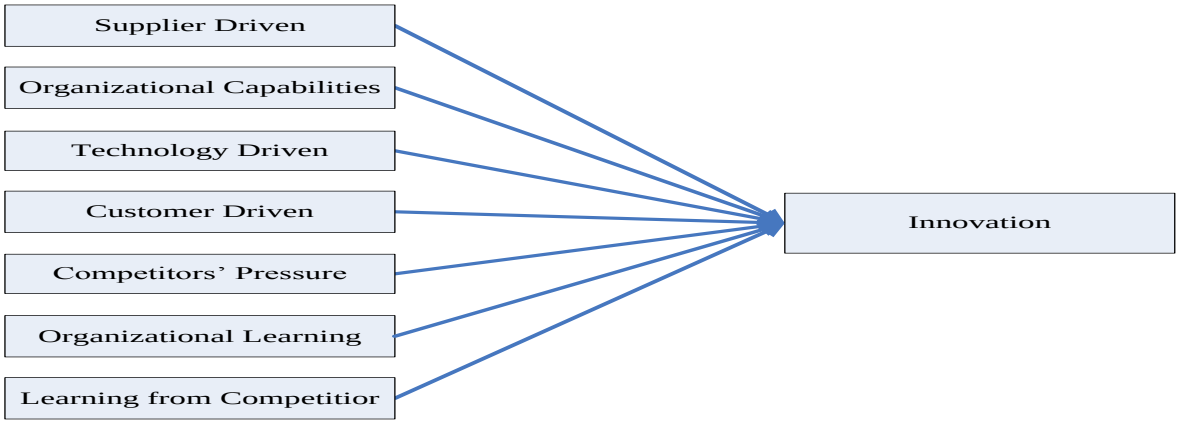


Figure 1. The conceptual model of the study

Table 1 shows the mean and standard deviation of all interesting factors. The factors that influenced the innovation include research and development, procurement department, sale departments, consumers, customers of your product, machine seller, pressure from competitors, impulse by the technology in the market, learning by studied from other companies, learning by studied from universities, learning by studied form public research institute, and organizational learning.

Based on the samples, learning by studied from other companies was the most influential factors to develop innovation at the average score of 3.7654 followed by research and development at the average score of 3.7284. The samples also suggest that pressure from competitors and impulse by the technology in the market are another two important factors which ranked as the third and forth representing the average score of 3.6296 and 3.5926, respectively. The consumer factor got the average score of 3.5185.

Table 1. Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ReDevelop	81	1.00	5.00	3.7284	1.14031
Procurement	81	1.00	5.00	3.3827	1.07898
SaleDepart	81	1.00	5.00	3.4691	1.32369
Consumers	81	1.00	5.00	3.5185	1.25610
CusAT	81	1.00	5.00	3.1975	1.21881
MacSeller	81	1.00	5.00	3.0247	1.45753
Competitiors	81	1.00	5.00	3.6296	.94133
Technology	81	1.00	5.00	3.5926	.73786
LearnComp	81	1.00	5.00	3.7654	.77897
LearnUni	81	1.00	5.00	2.6914	1.43738
LearnPublic	81	1.00	5.00	2.6420	1.21767
OrganLearn	81	1.00	5.00	3.3086	.81612

Table 2. Correlations between each factors

		ReDevelop	Procurement	SaleDepart	Consumers	ConAT	MacSeller	Pressure	Technology	Comp	Uni	Public	OrgLearn
ReDevelop	Pearson Correlation	1	.594**	.400**	.257*	.147	-.244*	-.281*	.030	.350**	-.120	-.053	.078
	Sig. (2-tailed)		.000	.000	.021	.190	.028	.011	.789	.001	.284	.639	.490
Procurement	Pearson Correlation	.594**	1	.730**	.221*	.322**	.042	-.166	.026	.465**	.117	.305**	.219*
	Sig. (2-tailed)	.000		.000	.048	.003	.712	.138	.821	.000	.297	.006	.049
SaleDepart	Pearson Correlation	.400**	.730**	1	.175	.399**	.124	-.110	-.045	.351**	.189	.323**	.154
	Sig. (2-tailed)	.000	.000		.118	.000	.272	.330	.690	.001	.092	.003	.171
Consumers	Pearson Correlation	.257*	.221*	.175	1	.496**	.102	.154	.231*	.228*	.124	.188	.183
	Sig. (2-tailed)	.021	.048	.118		.000	.364	.170	.038	.041	.269	.092	.101
ConAT	Pearson Correlation	.147	.322**	.399**	.496**	1	.335**	.130	-.048	.023	.192	.410**	.152
	Sig. (2-tailed)	.190	.003	.000	.000		.002	.248	.668	.838	.086	.000	.177
MacSeller	Pearson Correlation	-.244*	.042	.124	.102	.335**	1	.736**	.114	.159	.720**	.709**	.540**
	Sig. (2-tailed)	.028	.712	.272	.364	.002		.000	.311	.155	.000	.000	.000
Pressure	Pearson Correlation	-.281*	-.166	-.110	.154	.130	.736**	1	.446**	.170	.543**	.395**	.346**
	Sig. (2-tailed)	.011	.138	.330	.170	.248	.000		.000	.130	.000	.000	.002
Technology	Pearson Correlation	.030	.026	-.045	.231*	-.048	.114	.446**	1	.484**	.175	.086	.419**
	Sig. (2-tailed)	.789	.821	.690	.038	.668	.311	.000		.000	.119	.445	.000
LinkComp	Pearson Correlation	.350**	.465**	.351**	.228*	.023	.159	.170	.484**	1	.303**	.293**	.489**
	Sig. (2-tailed)	.001	.000	.001	.041	.838	.155	.130	.000		.006	.008	.000
LinkUni	Pearson Correlation	-.120	.117	.189	.124	.192	.720**	.543**	.175	.303**	1	.807**	.476**
	Sig. (2-tailed)	.284	.297	.092	.269	.086	.000	.000	.119	.006		.000	.000
LinkPublic	Pearson Correlation	-.053	.305**	.323**	.188	.410**	.709**	.395**	.086	.293**	.807**	1	.578**
	Sig. (2-tailed)	.639	.006	.003	.092	.000	.000	.000	.445	.008	.000		.000
OrganLearn	Pearson Correlation	.078	.219*	.154	.183	.152	.540**	.346**	.419**	.489**	.476**	.578**	1
	Sig. (2-tailed)	.490	.049	.171	.101	.177	.000	.002	.000	.000	.000	.000	

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

4.2 Data analysis

Most researchers use correlations to summarize the association between two scale variables. The correlation between two variables reflects the degree to which the variables are related. Generally, when discussing on correlations, it refers to Pearson's correlation coefficient. This value measures the strength of the linear relationship between variables. The correlation value coefficient can have values from -1 to +1. A correlation of 0 indicates that there is no linear relationship between the two variables. Correlations of -1 and +1 indicate that there is a perfect linear relationship in a negative or positive direction, respectively. This study performs correlation analysis in order to test the research hypotheses. If the hypotheses are correct, we should expect to find a positive correlation between factors and innovation.

Table 2 shows correlation result, which displays correlations among every factor specified. The second column shows the relationship between factors; which are research and development, procurement department, sale departments, consumers, customers of your product, machine seller, pressure from competitors, impulse by the technology in the market, learning by studied from other companies, learning by studied from universities, learning by studied form public research institute, and organizational learning. There are all positive correlation between these factors with the different strength. Another value to confirm the linear relationship is the value of sig. 0.000 or 0.001 which are statistically significant relationships between these desired factors. In this study, we are analyzing the factors by using the factor analysis technique to reduce the number of factors since there are many factors in this study. Table 3 shows the KMO (Kaiser-Meger-Olkin) value of 0.676, which indicated that these factors are related with each other significantly. Besides, the value of Bartlett's Test account for 555.234, which also shows that these factors are related together significantly with less than 0.05.

The principal component analysis (CPA) is used to analyze for combining each factors into group which relating to each factor significantly statistically. Table 4 shows the standardized factors which have four components extracting from 12 factors with 78.099 % of sample of total variance. When considering factor loading as shown in Table 5, it can be categorized four related groups including supplier driven, technological driven, customer driven, and organizational capabilities. The supplier driver consists of machine seller, learning by studied from universities and learning by studied from public research institute. The factors which are research and development, procurement department, and sale department that are categorized into organizational capabilities. Besides, technology as the third component factor. The fourth group is including consumers and customers. The rest of factors which not relate with other factors are consists of pressure of competitors, learning by studied from other companies and organizational learning. These are used to analyze in a single component.

Therefore, the multiple linear regressions in this study contain seven independent quantitative variables, which are supplier driven, technological driven, customer driven, organizational capabilities, pressure from competitors, learning b studied from other companies and organizational learning with the innovation as the dependent variable.

Table 6 illustrates the result from enrolling the independent variables into multiple linear regressions model with enter variables technique. The multiple regression model shows that the relationship between innovation and independent variables including supplier driven, technological driven, customer driven, organizational capabilities, pressure from competitors, learning b studied from other companies and organizational learning with the adjusted R-square 0.402 (40.2%). With regard to the Durbin-Watson value, the result value shows that the value 2.368 stands between the criteria value within the interval of 1.5 to 2.5. It can be concluded that the expected value of the residues are independent.

Table 3. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.676
Bartlett's Test of Sphericity	Approx. Chi-Square	555.234
	df	66
	Sig.	.000

Table 4 Total Variance Explained

Factor	Initial Eigen values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Var	Cumula-tive %	Total	% of Var	Cumula-tive %	Total	% of Var	Cumula-tive %
1	4.074	33.951	33.951	4.074	33.951	33.951	3.398	28.316	28.316
2	2.593	21.610	55.561	2.593	21.610	55.561	2.597	21.641	49.957
3	1.542	12.851	68.412	1.542	12.851	68.412	1.876	15.637	65.594
4	1.162	9.686	78.099	1.162	9.686	78.099	1.501	12.505	78.099
5	.597	4.976	83.075						
6	.506	4.213	87.288						
7	.461	3.841	91.129						
8	.330	2.753	93.882						
9	.299	2.489	96.371						
10	.231	1.925	98.296						
11	.114	.951	99.247						
12	.090	.753	100.000						

Extraction Method: Principal Component Analysis.

Table 5 Rotated Component Matrix[a]

	Component			
	1	2	3	4
ReDevelop	-.289	.722	.200	.163
Procurement	.100	.898	.067	.121
SaleDepart	.209	.811	-.103	.149
Consumers	.005	.124	.276	.858
ConAT	.301	.256	-.225	.784
MacSeller	.913	-.104	.051	.140
LearnUni	.875	.066	.124	.002
LearnPublic	.876	.248	.007	.128
Technology	.085	-.102	.887	.109
LearnComp	.203	.505	.696	-.061
Competitiors	.644	-.388	.380	.187
OrganLearn	.584	.195	.510	.001

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

The table 7 and table 8 shows the result of hypothesis testing relationship between innovation and independent variables including supplier driven, technological driven, customer driven, organizational capabilities, pressure form competitors, learning b studied from other companies and organizational learning with ANOVA and t-test. First, the result from ANOVA table indicated that the innovation determinant as dependent variable depended on at least one independent variable by statistical significance.

The second step is to test the relationship between innovation (constant) and each independent factor as a pair-wise comparison time by time. Table 8 illustrates the result from t-test value indicating that all of the five independent variables including supplier driven, technological driven, customer driven, organizational capabilities, pressure from competitors, learning by studied from other companies and organizational learning influenced the innovation determinant.

Table 6 Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.674[a]	.454	.402	.31578	2.368

a. Predictors: (Constant), Competitiors, LinkComp, CustomerDriven, OrganLearn, OrgCapabilities, SupplierDriven, TechnoDriven; b. Dependent Variable: Innovation

Table 7 ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.057	7	.865	8.678	.000[b]
	Residual	7.279	73	.100		
	Total	13.337	80			

b. Predictors: (Constant), Competitiors, LinkComp, CustomerDriven, OrganLearn, OrgCapabilities, SupplierDriven, TechnoDriven

Table 8 Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.865	.784		4.932	.000
	TechnoDriven	.244	.104	.597	2.339	.022
	OrganLearn	.166	.090	.332	1.855	.068
	CustomerDriven	.125	.038	.307	3.263	.002
	OrgCapabilities	.078	.059	.191	1.322	.190
	SupplierDriven	.054	.099	.133	.550	.584
	Competitiors	-.161	.088	-.371	-1.821	.073
	LinkComp	-.150	.106	-.286	-1.411	.162

a. Dependent Variable: Innovation

Surprisingly, the factors influencing innovation are only technological driven, customer driven, by statistical significance. The summarized linear regression model is as follows;

Innovation = 3.865 + 0.244 (Technological driven) + 0.125 (Customer driven)

4 Conclusion

In this study there has been a discussion on the innovation including characteristics and factors in the automotive industry in Bangkok, Thailand. The central objective is to investigate factors including supplier driven, technological driven, customer driven, organizational capabilities, pressure form competitors, learning by studied from other companies and organizational learning have an influence on innovation measurement within the context of automotive business in Bangkok of Thailand.

This study found that the automotive business is a technological-oriented business because the automotive parts have to porduce with the limited time and various shape to fulfill the need of the customers. The most significant factor that plays an important in the quality of the automotive industry is technology driven, especially in impulsing the new technology into the market. The data suggested that in order to achieve innovative activities, the firms must conduct some training activities and introduce the new production process using the new technology that relate to innovation. The better organization structure can be developed through management by issuing policies that encourage innovation; for example, innovation based on the acquisition or purchase new technology, the collaboration with suppliers, and the closely relationships with the desired customers. The firm's performance can be expected to improve by implementating technology. In particular, the market-related aspects such as market attractiveness, market responsiveness, and marketing strategies are also critical aspects for innovation.

In conclusion, the most significant factor that makes firms most successful is innovation. The important resource for implementing novel knowledge to enhance organization's ability is to develop innovative products as well as create new valuable services and competitiveness in an organization [4, 6, 7].

There is one major basic limitation of the study. The results refer to the fact that the percentage of respondent rate providing information is very low. This directly impacts to the result of the statistical data analysis. This leads to the unreliability of the summarizing result; for example, the tolerance values of colinearity statistics is all low and separately for each relevant criteria. This tolerance must close to 1.00. But the result is decreasing near 0.00. This is hardly to confirm that the multiple linear regression models is suitable for small sample size of the data.

Finally, in our study we employed the analysis of ANOVA and regression using static data, but if the causation was to be investigated the structural equation modeling methodology should be applied using longitudinal dynamic data [4].

References

[1] Amara, N., Landry, R., Becheikh, N., & Ouimet, M. Learning and novelty of innovation in established manufacturing SMEs. *Technovation*, 28(7):450-463, 2008.

[2] Armbruster, H., Bikfalvi, A., Kinkel. S., & Lay, G. Organizational innovation: The challenge of measuring non-technical innovation in large-scale surveys. *Technovation*, 28 (10):644-657, 2008.

[3] Bahinipati, B . K. & Deshmukh, S. G. Vertical collaboration in the semiconductor industry: A decision framework for supply chain relationships. *Computers & Industrial Engineering*, 62, 504–526,2012.

[4] Bahinipati, B. K., Kanda, A. & Deshmukh, S. G. Horizontal collaboration in semiconductor manufacturing industry supply chain: An evaluation of collaboration intensity index. *Computers & Industrial Engineering*, 57:880–895, 2009.

[5] Belton, V., & Stewart, T. R. Multiple criteria decision analysis: An integrated approach. Norwell, NA: Kluwer Academic Publishers, 2009.

[6]. Cao, M., & Zhang, Q. Supply chain collaborative advantage: A firm's perspective. International *Journal of Production Economics*, 128(1):358–367, 2010.

[7]. Cao, M., & Zhang, Q. Supply chain collaboration: Impact on collaborative advantage and firm performance. *Journal of Operations Management*, 29(3):163–180, 2011.

[8] Cousins, P. D. A conceptual model for managing long-term interorganisational relationships.

European *Journal of Purchasing and Supply Management*, 8(2):71–82, 2002.
[9] Feng, B., Fan, Z. P., & M, J. A method or partner selection of codevelopment alliances using individual and collaborative utilities. *International Journal of Production Economics*, 124:159-170, 2010.
[10] Feng, B., Fan, Z. P., & L, Y. A decision method for supplier selection in multi-service outsourcing. *International Journal of Production Economics*, 132:240-250, 2011.
[11] Koc & Ceylan, C. Factors impacting the innovative capacity in large-scale companies. *Technovation*, 27 (3):105-114, 2007.
[12] Laforet, S. Size, strategic, and market orientation affects on innovation. *Journal of Business Research*, 61(7):753-764., 2008.
[13] Mohamed, M. Z. Innovation implementations in Malaysian firms: Process, problems, critical success factors and working climate." *Technovation*, 15(6):375-385, 1995.
[14] Nijssen, E. J., Hillebrand, B., Vermeulen, A. M. & Kemp, G. M. Exploring product and service innovation similarities and differences. International *Journal of Research in Marketing*, 23(3):241-251, 2006.
[15] Organisation for Economic Co-operation and Development-OECD. *Oslo Manual: Proposed Guildlines for Collecting and Interpreting Technological Innovation Data*. Paris, 2005.
[16] Rosenbusch, N., Brinckmann, J., & Bausch, A. Is innovation always beneficial? A meta-analysis of the relationship between innovation and performance in SMEs." *Journal of Business Venturing*, (In press).
[17] Smith, D. Exploring innovation, UK: *Mc Graw-Hill*, 24-45, 2010.
[18] Soosay, C. A., Hyland, P. W. & Ferrer, M. Supply chain collaboration: capabilities for continuous innovation. *Supply Chain Management: An International Journal*, 13(2), 160-169, 2008.
[19] Thitisomboon, T., Visanvetchakij, T., Chalermchutidet, C., Phongarjarn, E. & Jeenanunta, C. Supply Chain Collaboration in Automotive Industry: A Case Study of Toyota (Thailand). *Presented at the AIB South East Asia Regional Conference*, Hong Kong, CH, 2009.
[20] Wan, D., Ong, C, H. & Lee, F. Determinants of firm innovation in Singapore. *Technovation*, 25(3), 261-268, 2005.
[21] Zeng, S. X., Xie, X. M. & Tam, C. M. Relationship between cooperation networks and innovation performance of SMEs. *Technovation*, 30, 181-194, 2010.

Developing an Appropriate Data Collection Process on Traffic Accident in Thailand

Nitipon Tansakul **Suthathip Suanmali*** **Veeris Ammarapala**

Transportation Research Center, Sirindhorn International Institute of Technology
Thammasat University, Pathumthani 12000 Thailand
*Corresponding author: ssuanma@gmail.com

Abstract

Traffic accident is the major problem for many countries, resulting in death and damaged for both people and economies. Their are three main factors that cause an accident; they are road and environment, vehicle, and human. In order to identify the critical reason of traffic accidents, the data collection process must be efficient and easy to understand. However, the accuracy is the most important factor to identify the cause of an accident. Therefore, the objective of this paper is to develop an efficient accident report form to identify the problem of road accident in Thailand. The expected result is to identify the specific causes of accidents and enable the policy makers to seek appropriate policies and strategies that enhance road safety among highways in Thailand.

Keywords: Accident rate, Acident Report Form, Index of item-objective congruence (IOC), Road Safety

1 Introduction

An accident occurred on national highways resulting in death, injury, property damaged, and even an economic loss. Therefore, many government agencies have been studying on the road accidents in order to reduce the number of deaths and injuries. The study of Thailand Accident Research Center (TARC) in 2009, presents three cofactors that cause an accident; they are human, vehicle, and environmental factors (Figure 1). The highest percentage of accidents occurs because of human following by environment and vehicle factors respectively. Although, it can be typically concluded that the accident is caused by the failure of motorists who could not safely drive and control a car. However, road and environment can cause an accident as well, both from direct and indirect perspectives. Direct cause includes damage surface, dangerous intersection and curve. Indirect cause includes roadside asset such as traffic sign, electricity post, street light and equipments located on the sidewalk (TARC, 2009). The study of TARC has found that whenever the car is out of the main path, it would have collided with these objects and caused death or serious injury.

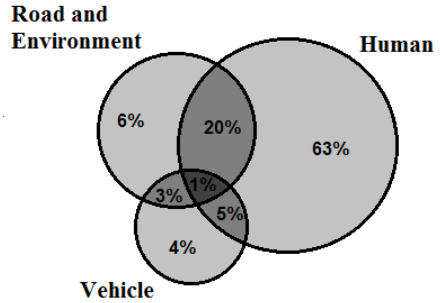


Figure 1. Elements of the Accident

Statistical analysis is an important tool to analyze traffic accident data. The Bureau of Highways Safety, Department of Highways (DOH) has the objective to develop strategy plans for transportation and safety in both short term and long term. Thereby, it is necessary to study and explore the feasibility analysis of project planning to raise safety standards. The traffic accidental data on national highways is usually collected by the field district officers. All accidental data is submitted directly to Bureau of Highway Safety. After that, the accident statisticians will summarize and present the results in the annual report. Figure 2 shows the number of accidents, injuries, and deaths between 2007 and 2012; the number of accidents decreases 19 percent, and the number of injuries decreases 13 percent, but number of death is almost the same.

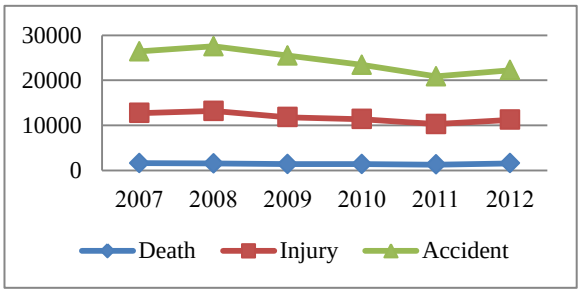


Figure 2. Number of Accident, Injury, and Death in Thailand from 2007 to 2012

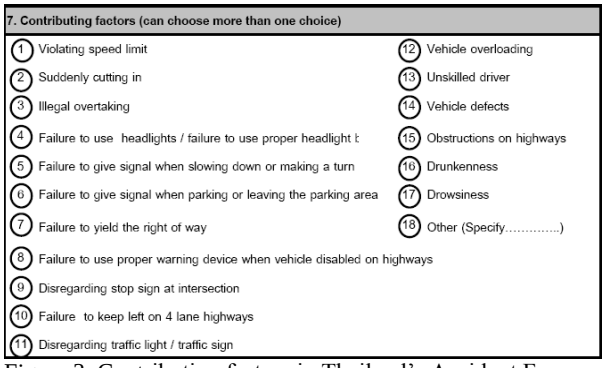
DOH classifies the causes of accident into many factors according to Table 1 below. Observe that violating speed limit is the most factors that cause the accidents. In 2011, there are 8,287 accidental causes occurred on the highways nationwide and 78.3 percent is caused by violating the speed limit following by suddenly cutting in (8.07 percent) and illegal overtaking (1.55%).

Table 1. Result from Current Accident Report Form in Thailand

Accident Causes	Pedes-trian	On path	Off path on straight	Off path on curve	Other	Total	Percent-age
1. Direct human causes	302	7862	875	410	130	9579	90.31%
Violating speed limit	140	6878	834	349	86	8287	78.13%
Suddenly cutting in	102	646	24	42	42	856	8.07%
Illegal overtaking		150	3	9	2	164	1.55%
Failure to headlights / Use proper headlight beam		7				7	0.07%
Failure to give signals when slowing down or marking a turn	3	44		4		51	0.48%
Failure to give signals when parking or having the parking area		17				17	0.16%
Failure to yield the right of way	12	4				16	0.15%
Failure to use proper warning device when vehicle disabled on highways	1	6	1	1		9	0.08%
Disregarding stop sign at intersection	7	1				8	0.08%
Failure to keep left on 4 lane highways		6				6	0.06%
Disregarding traffic light / traffic sign	36	35	2	3		76	0.72%
Unskilled driver	1	68	11	2		82	0.77%
2. Indirect human causes	10	572	72	25	5	684	6.45%
Drunkenness	6	195	28	20	5	254	2.39%
Drowsiness	4	377	44	5		430	4.05%
3. Vehicle causes	3	169	52	6	1	231	2.18%
Vehicle overloading	1	15	11	3		30	0.28%
Vehicle defects	2	154	41	3	1	201	1.89%
4. Other		86	10	17		113	1.07%
Total	315	8689	1009	458	136	10607	100.00%

1.1 Problem Statement

Although the traffic accident statistics in Thailand are collected every year, the data are not presented with reasonable results for developing the effective policy to reduce the accident rate. In the present, violating the speed limit has the highest proportion of causing the accidents is number one cause of accident. It is undeniable that many drivers are driving above the speed limit when accidents occur, but it is not reasonable to solely conclude that high speed driving is the crucial cause of each accident. In addition, when looking the at the accident report form currently used by the officers, it is not surprising to see such a conclusion indicating that most accidents are caused by high speed driving. It is believed that the accident report form currently used is not effective. The current form offers limited choices of contributing factors that cause an accident (Figure. 3). In addition, most of the time accidents occur with more than one contributing factors (Austroads, 1994). However, according to Table 1, violating speed limit is the most serious problem that causes traffic accidents (8,287 times in 2011) with a huge gap far from the second one, suddenly cutting (856 times in 2011). Therefore, it is necessary to clearly define the crucial causes of traffic accidents. If they are not correctly determined, authorities will not be able to apply the correct methods or measures in order to effectively lower the number of accidents, deaths and injuries. Thereby, it is necessary to develop an effective accident report form in order to collect meaningful accidental data. Then the analyzed data will identify the specific causes of accidents and enable the policy makers to seek appropriate policies and strategies that enhance road safety.


7. Contributing factors (can choose more than one choice)

1 Violating speed limit
2 Suddenly cutting in
3 Illegal overtaking
4 Failure to use headlights / failure to use proper headlight b
5 Failure to give signal when slowing down or making a turn
6 Failure to give signal when parking or leaving the parking area
7 Failure to yield the right of way
8 Failure to use proper warning device when vehicle disabled on highways
9 Disregarding stop sign at intersection
10 Failure to keep left on 4 lane highways
11 Disregarding traffic light / traffic sign
12 Vehicle overloading
13 Unskilled driver
14 Vehicle defects
15 Obstructions on highways
16 Drunkenness
17 Drowsiness
18 Other (Specify.....)

Figure 3. Contributing factors in Thailand's Accident Form

1.2 Research Objective

The objective of this research is to review existing procedure of collecting accidental data. A new accidental report form is proposed for improving the data collection procedures. The data gathered from this newly developed form will enable authorities to completely analyze for crucial factors that are contributing to each accident. It is planned to study the process for collecting data in other countries to better identify the causes of traffic accidents and a root of the contributed factors in Thailand. Once the data are better collected, and authorities are able to analyze for a meaningful conclusion of each accident, they will be able to develop an effective technique and policy in order to prevent any accident effectively in the future.

2 Literature Review

Accident statistics data are effective information for developing the policy to prevent the traffic accident. Many countries have developed their accident report form to collect a specific data. Then those data were used for researching and developing innovation, policies, tools, and technique to prevent an accident and reduce injury and death.

2.1 Data Collection and Result

An efficient data collection process will enhance an accuracy of the study. The report of New Zealand emphasis on the multi-factors that cause an accident. Thus, the report presents the comparable relative risk of accident for drunk-driving and speeding driving. At the speed above 60 kilometers per hour zone, increasing 5kph will increases the risk of a crash by about the same as an increase in blood alcohol concentration (BAC) of 50 mg/100 ml (LTSA, 2000). The statistical data in New Zealand, for the years 2009 to 2011, is shown in Figure 4 indicating that speeding contributes to a road accident about 30 percent. The combination of speeding and alcohol/drugs driving contributes to 17 percent of fatal accidents and 13 percent of fatal accidents occur from high speed driving only.

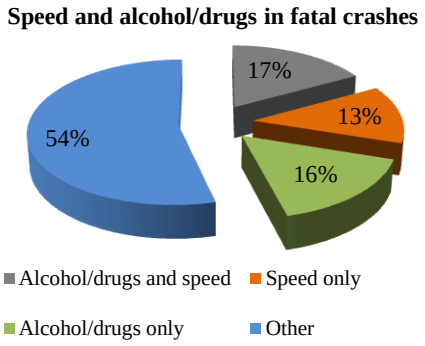


Figure 4. Speed and alcohol/drugs in fatal crashes in New Zealand years 2009 to 2011

Over speeding is one of the major causes in a traffic accident, but other factors such as drunkenness and drowsiness contributes to traffic accident as well. The purpose of setting up a speed limit on each highway is to improve traffic safety (A. Dutta and D. A. Noyce). Speed limits have an effect to reduce probability of crashes and violence of the collision. In Thailand, over speeding has caused 78 percent of total accidents. From Table 2, speeding has caused only 17.2 percent of crashes in New South Wales. However, 40.8 percent of crashes involved by the speed are a fatal crash. It means speeding can increase violence of crash.

Table 2. Casualties, Speeding Involvement in Crash, Degree of Casualty

Speeding involved in crash	Degree of crash			
	Fatal Crash	Injury crash	Non-casualty crash	total crashes
Yes	137	3393	3866	7396
No or unknown	199	16512	18846	35557
Total	336	19905	22712	42953
Percentage	40.8%	17.0%	17.0%	17.2%

The United State of American has many studies on the traffic accidents. The crash event data are collected and contributed to many categories. National Highway Traffic Safety Administration or NHTSA has a mission of reducing crash-related injuries and fatalities while ensuring the highest standards of safety on the nation's roadways. NHTSA has conducted the National Motor Vehicle Crash Causation Survey (NMVCCS) of crashes with focus on the factors related to pre-crash events involving light passenger vehicles. They have completed a nationwide survey investigated during a two-and-a half-year period (July 3, 2005, to December 31, 2007) by using survey and case study method in order to determine the crash event.

Since, each state of the United State of America has developed their *Crash Report Forms* independently. NHTSA supports the use of uniform police crash report by setting up a standard of collecting data in order to encourage the field officer to understand the pre-crash events and collect the crash data in the same standard over 50 states. *Manual On Classification Of Motor Vehicle Traffic Accidents* (ANSI D-16) is used to provide a common language for a field officer to collect the information on traffic crash, and *Data Element Dictionary For Traffic Records Systems* (ANSI D-20) is used to provide a common set of data element coding instructions as these relate to traffic safety, driver licensing and vehicle registration.

The result information about these critical reasons can be attributed to a driver, vehicle, or environment. Driver-related critical reasons, broadly classified into recognition errors, decision errors, performance errors, and nonperformance errors. Each error type consists of many sub factors as shown in the table below. The highest percentage of the pre - crash event is inadequate surveillance which is about 20 percent, when too fast for conditions and too fast for curves is only 8.4 percent and 4.9 percent respectively.

Table 3. Critical Reason for Critical Pre-Crash Event Related to Human Error

Critical Reason for Critical Pre-Crash Event	Number of Crashes		Weighted Percentage
	Unweighted	Weighted	
Recognition error			
Inadequate surveillance	1080	414626	20.3%
Internal distraction	482	218548	10.7%
External distraction	229	77496	3.8%
Inattention (i.e., daydreaming, etc.)	194	65712	3.2%

Other/unknown recognition error	109	51926	2.5%
Subtotal	2094	828308	40.6%
Decision error			
Too fast for conditions	348	171604	8.4%
Too fast for curve	181	100713	4.9%
False assumption of other's action	260	92583	4.5%
Illegal maneuver	232	78112	3.8%
Misjudgment of gap or other's speed	212	65221	3.2%
Following too closely	85	30452	1.5%
Aggressive driving behavior	99	31026	1.5%
Other/unknown decision error	335	125805	6.2%
Subtotal	1752	695516	34.1%
Performance error			
Overcompensation	211	100090	4.9%
Poor directional control	249	95165	4.7%
Other/unknown performance error	30	7751	0.4%
Panic/freezing	20	7137	0.3%
Subtotal	510	210143	10.3%
Non-performance error			
Sleep, actually asleep	160	65141	3.2%
Heart attack or other physical impairment	133	48822	2.4%
Other/unknown critical nonperformance	76	31881	1.6%
Subtotal	369	145844	7.1%
Other/unknown driver error	371	162132	7.9%
Total	5096	2041943	100.0%

Vehicles are another important reason of causing a traffic accident, a damaged part or component within the vehicle which is not well maintained can cause a fatal accident. NHTSA also collects the information about critical cause related to the vehicles. In Table 4, the critical reason for pre-crash event related to vehicles, the most frequently occurring was tire failure or degradation/wheel failure, which is about 43 percent of the crashes related to vehicle error. Brake failure/degradation is assigned to 25 percent of total crashes related to vehicle error following by steering/suspension/transmission/engine failure that assigned to 10.5 percent of the crashes.

Table 4. Critical Reason for Critical Pre-Crash Event Related to Vehicle Error

Critical Reason for Critical Pre-Crash Event	Number of Crashes		Weighted Percentage
	Unweighted	Weighted	
Tires failed or degraded/wheels failed	56	19320	43.3%
Brakes failed/degraded	39	11144	25.0%
Other vehicle failure/deficiency	17	9298	20.8%
Steering/suspension/transmission/engine failed	16	4669	10.5%
Unknown	2	212	0.5%
Total	130	44643	100%

In summary, there are a lot of factors contribute to a traffic accident. It was adopted according to the characteristic of crash in each area. Some factors can be adjusted such as in cool weather countries, snowing is a problem that typically causes the accident. However, the field office

also is an important factor as he or she is the one that classifies the critical reason of the accident. Moreover, the study of TARC presented that an accident mainly relies on three cofactors; which are human, vehicle, and environmental factors. Therefore, Thailand's accident report form should consist of those three factors. Since the accident can occur from more than one contributing factors then it is important to train the field officer to classify main factors and co-factors in each accident.

2.2 Reliability of the Instruments Used in the Research

In order to check the validity of the form before being used to collect the actual data. The index of item-objective congruence (IOC) concept is used for evaluating the content validity on measure specific objective at the development stage, develop by Rovinelli and Hambleton in 1977. This concept is a process that a group of content experts individually evaluates the list of items based on the degree which they can measure specific objective to verify the accuracy of the contents (Ronna C. Turner & Laurie Carlson 2003). Each item will be evaluated by the expert, rating of 1 if expert definite feeling that an item is a measure of an objective, 0 if expert undecided about whether the item is a measure of an objective, or -1 if expert definite feeling that an item is not a measure of an objective. It is usually considered three to seven experts to conclude the consistency index between item and objective. Then the index of item objective congruence will be calculated by the formula below;

$$IOC = \frac{\sum R}{N}$$

Where;

IOC = Index of Item Objective Congruence

ΣR = Total score from expert

N = Number of expert

After that, the result of each item will range from -1 to 1, the cutoff score of this index would be 0.5 or 50 percent. Then the good items with IOC index more than or equal to 0.5 are accepted and bad items with IOC index less than 0.5 should be revised or eliminated.

3 Methodology

The objective of this study is to develop an effective accident report form to better collect a statistical data to identify the real problem that causes an accident. The form should be easy to understand for a field officer and cover as many causes of an accident as possible. The contributing factor in the form must be appropriate to the characteristics of a traffic accident in Thailand. Most of the time, accident occur with multi factors so this form will allow any officer to choose more than one choice. Then it can be used to identify both main factor and catalyst factor of an accident. The traditional accident report form is able to collect effective information for many topics, but the contributing factors of an accident (Topic # 7) are not clearly defined. So, it can persuade field officers to misunderstand then they may fill in the form with less recognition. In order to efficiently collect the data, report form of an accident should clearly define each factor and divided into a group of contributing factors as shown in Table 5.

Table 5. New Contributing Factors of an Accident

Contributing Factors of An Accident	USA (NHTSA)	Thai (TARC)	USA (Florida)	Developed
Recognition error				
Inadequate surveillance	✓		✓	✓
Internal distraction (i.e., telephone, etc.)	✓			✓
External distraction (i.e., stare at somewhere, etc.)	✓			✓
Inattention (i.e., daydreaming, etc.)	✓			✓
Failure to use headlight		✓		✓
Failure to give signal to other cars		✓		✓
Disregarding traffic light/traffic sign		✓	✓	✓
Decision error				

Suddenly cutting in		✓	✓	✓
Too fast for conditions	✓	✓	✓	✓
Too fast for curve	✓	✓	✓	✓
False assumption of other's action	✓		✓	✓
Misjudgment of gap or other's speed	✓			✓
Following too closely	✓		✓	✓
Aggressive driving behavior	✓			✓
Performance error				
Overreact	✓			✓
Poor directional control	✓			✓
Panic/freezing	✓			✓
Non-performance error				
Drunkenness	✓	✓	✓	✓
Drowsiness	✓	✓		✓
Physical impairment	✓			✓
Vehicle causes				
Tire failed or degraded	✓		✓	✓
Brakes failed	✓		✓	✓
Steering/suspension/transmission/engine failed	✓		✓	✓
Overloading		✓	✓	✓

This form is based on the result data collection of pre-crash event developed by National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation, adjusted with the traditional report form of The Bureau of Highways Safety, Thailand Department of Highways (DOH) to cover the characteristic of traffic in Thailand. It consists of two important causes of an accident, which are human related factors (recognition error, decision error, performance error, and nonperformance error) and vehicle elated factors. In additional, the field researches should be trained to identify, collect, and document of the relevant information in order to prevent the mislead result (NHTSA, 2008).

However, before the new accident report form will be used for collecting the accident data in Thailand. It is necessary to check the validity of the developed form. IOC is employed to measure the validity and justify the ability of each item for measuring the objective. Three experts in related fields are judging the appropriation of each contributing factor of an accident, the measurement ranging from -1 to 1. Then, the expert can give suggestion to improve the compatibility of an item and objective. After the data were collected from three experts, those data will be calculated to find IOC index. If the result value of IOC index in each factor is higher than 0.5, it mean the factor in this form are accepted and the form can measure the traffic accident in Thailand.

4 Conclusion

The data collection is an important process for any study, this study will benefit to an accident analysis in Thailand. In the present, violating the speed limit is highly influential in an accident in Thailand. This study expected to see more reasonable cause of a traffic accident by breaking down the causes of an accident into many related factors in an accident report form. The form will better contribute the critical factor that related to crash event and enable a field officer to better identify a critical reason of the traffic accident.

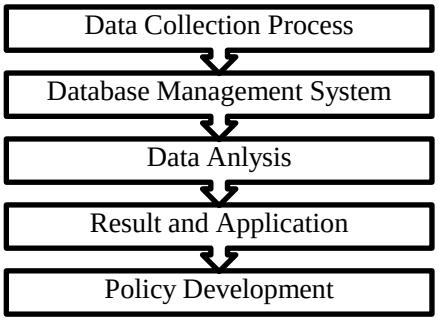


Figure 5. Data Collection and its Application

Figure 5 shows the process of policy development, the data collection process is the input

information for the any analysis. If the input is ineffective, then the result will be incorrect, and it will mislead the policy maker. Finally, the real problem will not be solved. Therefore, the improvement of the data collection process of related factors that cause an accident is necessary in order to reduce the number of accidents, injury, and death rate from traffic travel. Moreover, policy maker can efficiently develop a proper strategy to reduce crash or suggest any equipment and technique to improve road safety. Since new accident report form can better identify the root of traffic accident and develop the policies. In addition, these data can be used to find the relationship of accident with other factors such as gender, age, road surface, and etc. For instance, if most of the accidents occur on wet concrete surface. It should be considered to increase the friction of the road surface, high friction surfaces are consider to use in order to restore or enhancing the skid resistance of roadway.

However, if driving too fast for condition is the most influence problem in Thailand. The government should study to identify this problem. Then the result can be suggested as the following;

- Increasing punishments: If driver driving by aggressive speed without considering the safety of others, increasing the penalty of violating speed limit and number of tracking camera will increase the awareness of drivers. If a driver realizes high punishment, the rate of violating the law will possibly reduce.
- Improve road safety: If high speed driving is usually occur and catalyze any traffic accident. An Improvement of road safety should be considered. Nowadays, technology for road safety was invented for reducing an accident rate. For example, in order to prevent roadway departure crashes. Rumble Strips are invented, the noise and vibration produced by rumble strips will alert drivers when they about to leave their travel way. At the curve, rumble strips can alert the driver to reduce speed and attended to the situation ahead.

In addition, many traffic problems occur with the carelessness of drivers. Therefore, it can prevent if everyone recognizes safety and driving with responsibility. They are many campaigns try to motivated about this for preventing and increasing an awareness of the driver. The result of this data collection form can identify the group of the driver who cause more accident and the critical reason of the accident. Then the campaign can be adjusted to complete with the characteristic of those drivers.

References

[1] Traffic Accident on National Highways 2011. *Statistical Information Group, Bureau of Highway Safety*, Department of Highways, Ministry of Transport, Thailand, 2011.

[2] Thailand Accident Research Center. *To develop and Disseminate knowledge on road safety*, Final Report 2009.

[3] National Motor Vehicle Crash Causation Survey. *National Highway Traffic Safety Administration (NHTSA),* U.S. Department of Transportation 2008.

[4] Road Traffic Crashes in New South Wales the Centre for Road Safety. Transport for NSW Statistical Statement for the year ended 2011.

[5] Speeding. *Crash Statistics for The Year Ended 31 December 2011*, Crash Factsheet Ministry of Transport 2012.

[6] Down with Speed. *A Review of the Literature, and the Impact of Speed on New Zealanders*, Accident Compensation Corporation (ACC) and Land Transport Safety Authority (LTSA) 2000.

[7] Road Safety Audit. *AUSTROADS Publication*, Sydney, Australia, ISBN: 085588455X 1994.

[8] Arup Dutta and David A. Noyce. Impact of Raising Speed Limits on Traffic Safety. *Traffic Operations and Safety Laboratory* 2007.

[9] Rovinelli, R. J.,&Hambleton, R. K. On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2: 49–60, 1977.

[10] Ronna C. Turner & Laurie Carlson. Indexes of Item-Objective Congruence for Multidimensional Items, *International Journal of Testing*, 3(2): 163-171, 2003.

Acknowledgment

The authors would like to acknowledge Transportation Research Center (TREC) for providing the information and financial support to attend this conference

Innovation of Automobile Dealers in Hokkaido

Tomomi Kaneko† **Wataru Hase†** **Setsuko Takahashi†** **Ryoju Hamada‡**

†Hokkaido Automotive Engineering College, Japan
‡Tohoku University, Japan
kaneko@haec.ac.jp

Abstract

On December 29, 1989, the Nikkei 225 Stock Average closed at 38,915.87, the highest value of the index since records were first taken in 1949. The sentiment of the Japanese economy until the end of the 1980s was extremely good. Japanese people purchased many goods. Actually, domestic sales of new vehicles in 1990 were about 7.7 million [1]. However, after 1990, Japanese people continued to suffer through the long period of social and economic stagnation which came to be known as Japan's "Lost Two Decades." Domestic sales of new vehicles in 2011 decreased to about 4.2 million units [1], approximately half of the 1990's sales. Generally speaking, the profit of automobile dealers depends on new vehicles sales. Because of declining new vehicle sales, Japanese automobile dealers needed to make innovations related to their business model.

In this report, the authors describe the present situation of the automotive industry in Japan and innovations of automobile dealers in Hokkaido.

Keywords: Automobile dealer, New vehicle Sales, Automobile mechanics, Service after sales, Innovation

1 Outline of the Japanese Automotive Industry

1.1 Present situation of the automotive industry

After World War II, the Japanese automotive industry developed in tandem with economic growth. It became a key industry in Japan. Nowadays Japanese vehicles are extremely popular and widely recognized as reliable vehicles throughout the world. Actually, in 1999, the author rode a Japanese pick-up truck at a petroleum site in Saudi Arabia.

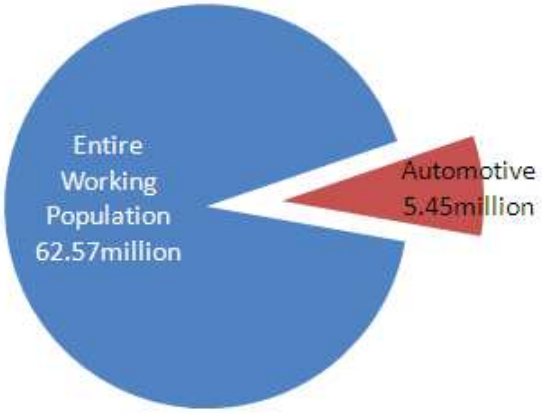


Figure 1. Automotive working population against entire working population in 2010 [2].

Figure 1 shows that the percentage of the population engaged in automotive industry in Japan is about 8.7% of the entire working population in 2010. In other words, one of 11.5 workers is engaged in the automotive industry. They engaged engineers, mechanics, sales people, drivers, and so on.

Figure 2 shows that the percentage of automotive shipments in Japan was about 16.4% of all manufacturing industry shipments in 2010. Many Japanese Automobile manufacturers produce and sell their brand vehicles all over the world, for example, Thailand, U.S.A., and China.

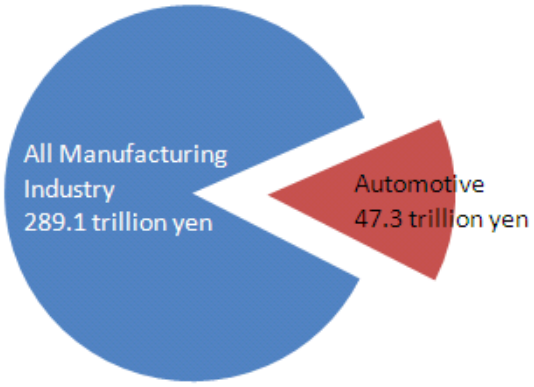


Figure 2. Automotive shipment value against total industry shipments in 2010 [3].

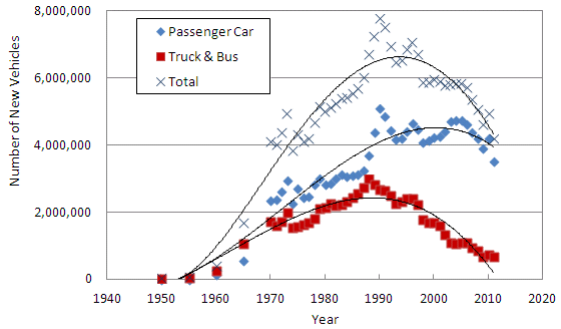


Figure 3. New vehicle sales in Japan [1].

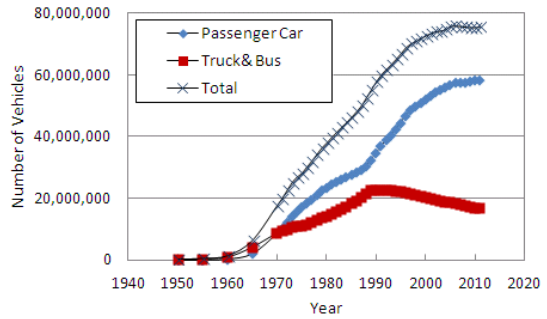


Figure 4. Registered vehicles in Japan [1].

Japanese vehicles have 85.9% market share in Thailand [1].

Figure 3 shows that sales of new vehicles reached a peak of 7.7 million in 1990. Thereafter, new vehicles sales decreased. In 2011, total new vehicle sales were almost equal to those of 1970. The breakdown of new vehicles sales means that between 4–5 million passenger cars have been sold annually from 1990. However, trucks and buses show a peak at around 1990 during the same period at the peak of Japan's "bubble economy".

Figure 4 shows that total registered vehicles hit a peak of 75.5 million vehicles around 2008. Numbers of truck and buses registered in Japan showed a peak around 1990. Subsequently, they decreased gradually. The reason for this trend is the stagnation of Japanese economic conditions. As for passenger cars, they have increased, partially because lifestyles of Japanese families have changed. Parents in Japan need passenger cars to take children to various after-school and cultural activities, private tutoring schools, and so on.

Figure 5 shows that the passenger car ownership rate per family in Japan was over 1.000 in 1996. Nevertheless, family units that include parents and children show almost no change from 1980. During the twenty-first century, the rate stayed between 1.075 to 1.112. The author infers that most families have already purchased passenger cars.

Figure 6 shows that the service life of vehicles is increasing every year. Japanese vehicles are well-known to be more or less trouble-free and can be used for longer than 13 years or for more than 100,000 km. Because of their high quality and durability, almost all Japanese people purchase Japanese vehicles. The market share of Japanese automobile manufacturers in Japan is 93%.

Figure 7 shows that Japanese GDP increased along with the number of registered vehicles in Japan until 2000. However, after 2000, Japanese GDP and registered vehicles did not increase.

Figure 8 shows that the peak year of population older than 15 years is the same period as the peak of registered vehicles in Japan. These trends indicate that the number of registered vehicles might decrease gradually from now.

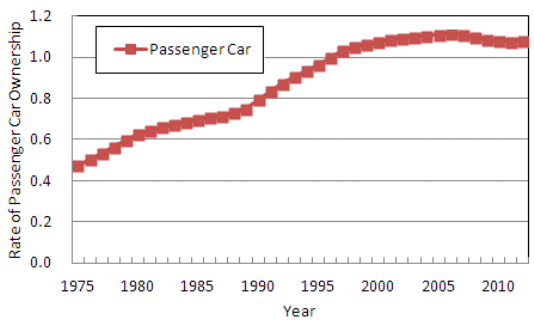


Figure 5. Passenger car ownership rate in Japan [6].

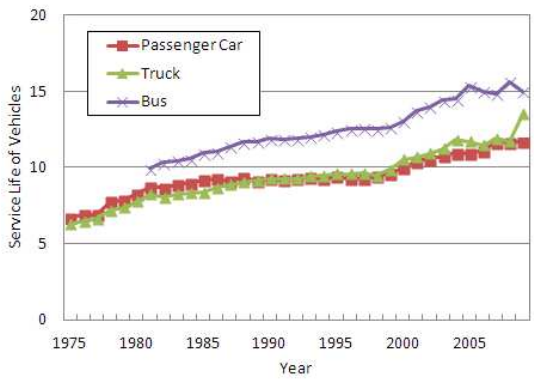


Figure 6. Service life of vehicles [1].

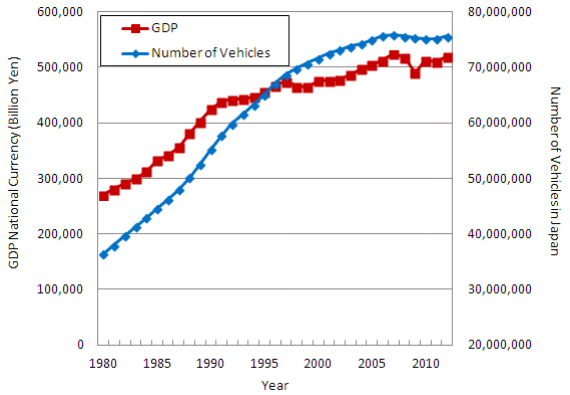


Figure 7. Japanese GDP (national currency) and number of vehicles in Japan [1], [4].

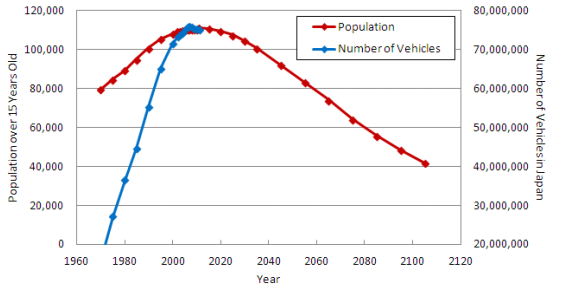


Figure 8. Population older than 15 years and the number of vehicles in Japan [1], [5].

1.2 Innovations of automobile dealers in Hokkaido

The discussion presented above implies that the number of registered vehicles in Japan might decrease and suggests that it will be increasingly difficult to sell new vehicles. Therefore, Japanese automobile dealers must change their business strategy. It is considered that they have three options:

(1) Sell various vehicles that the customer requests.
(2) Get a new owner.
(3) Change the business strategy without dependence on new vehicles sales.

They cannot choose option 1 because they are just one of an affiliated group of automobile manufacturers that include Toyota, Honda, and Nissan. They cannot participate in determining the strategy of new vehicles' release. They also cannot choose option 2 because the discussion presented above demonstrates that potential passenger car owners have owned cars. Automobile dealers are compelled to choose option 3.

Figure 9 shows that the proportion of new passenger car sales is decreasing and, on contract, the proportion of service is increasing for automobile dealers in Hokkaido.

Figure 10 shows that the sales composition of new passenger cars in 2003 closely resembles that of 2008. However, the profit composition of new passenger cars shows a 4.2% decrease. That

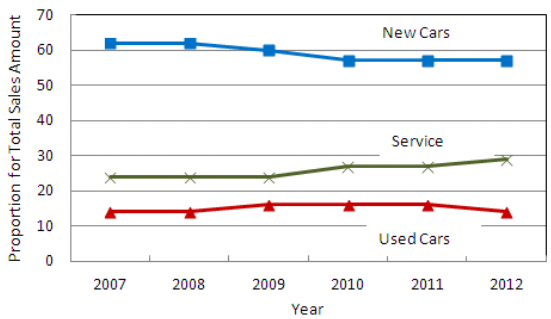


Figure 9. Changes in the proportion of total sales amount.

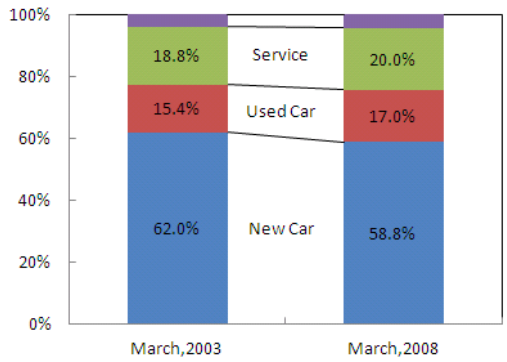


Figure 10. Sales compositions of automobile dealers in 2003 and 2008.

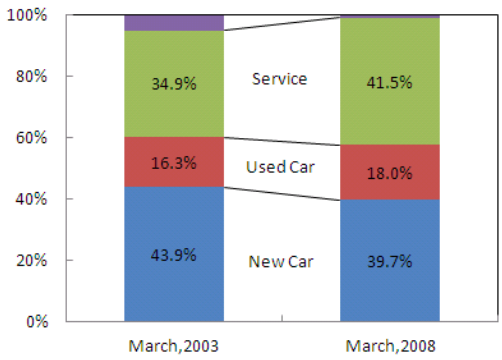


Figure 11. Profit composition of an automobile dealer in 2003 and 2008.

of service is a 6.6% increase, as shown in Figure 11. This result indicates that growth of the service branch is extremely important and that automobile dealers emphasize service provision because they have difficulty selling new passenger cars. They expect some customers to repurchase new passenger cars through the relation formed through a service branch.

2 Japanese Automobile Maintenance System

2.1 Automobile inspection system in Japan

Japan has some strict automobile inspection programs.

(1) Automobile safety inspection

Fundamentally, a passenger car that is driven in Japan must be inspected every two years. This is a mandatory inspection. If a passenger car is not confirmed with automobile safety inspection, it loses its registration and cannot be used.

(2) Periodic check (every 6 months and every 12 months).

A user of a passenger car should perform periodic checks. This is voluntary check.

The above two programs must be carried out by certified automobile mechanics.

2.2 Qualification system for automobile mechanics

The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) establishes qualification system for automobile mechanics as follows.

- First-class automobile mechanics
- Second-class automobile mechanics
- Third-class automobile mechanics

Actually, second-class qualification is sufficient to work at a maintenance shop. To become a qualified automobile mechanic, a paper examination and a practical examination must be passed.

2.3 Professional education for automobile mechanics

The most popular way to become an automobile mechanic is to enroll at an automotive engineering college or technical college. After finishing the two-year curriculum, graduates must pass a qualification test for automobile mechanics.

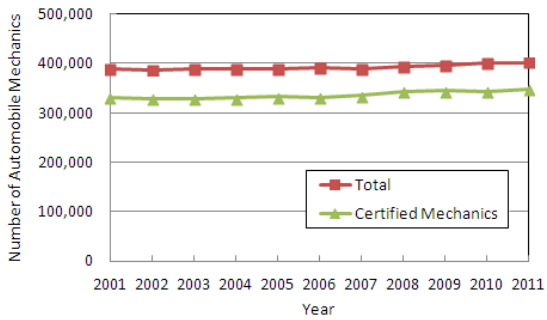


Figure 12. Automobile mechanics during 2001–2011 [7] [8].

Figure 12 shows that there has been no change of the number of automobile mechanics at maintenance shops. The reason for this trend is that the number of these colleges' students has been decreasing because automobile mechanics seem not to have good jobs. Moreover, young people have lost interest in cars. In fact, automobile dealers must confront a persistent shortage of automobile mechanics.

2.4 To acquire automobile mechanic skills

A decade ago, automobile mechanics were requested only to pass the qualification examination. While working at an automobile maintenance shop, they learned maintenance skills through on-the-job training. They worked just as automobile mechanics, that is to say, they did not explain the causes of trouble to the customers. Sales people and customer service employees at the shop explained matters to customers.

The author used to be a Vice Manager of the Career Development Division in Hokkaido Automotive Engineering College and had discussions with people in charge of recruitment officers of automobile dealers. They reported that today's automobile mechanics must not only have appropriate qualifications but also maintenance skills and communication skills. They must learn practical maintenance skills at college because senior mechanics at a maintenance shop are too busy to teach maintenance skills to young mechanics. Moreover, they must explain the cause of trouble to the customers personally. Their two skills bring more profits to automobile dealers.

3 Conclusions and Future Work

3.1 Conclusions

This paper presents the present situation of automotive industry in Japan and study of a business strategy of automobile dealers in Hokkaido.

(1) The automotive industry is extremely important in Japan.
(2) An innovation of automobile dealers is to increase the share of services against total profit in accordance with decreasing new passenger car sales. Therefore, they need more automobile mechanics, but they cannot hire a suitable number of automobile mechanics.
(3) In the past, automobile mechanics were requested to have a qualification of second-class automobile mechanics. However, to accommodate the innovative strategies of automobile dealers, they are requested not only to have that qualification but also to have practical maintenance skills and skills of communication with customers.

3.2 Future work

In our future work, we would like to investigate Japanese automobile dealers further and conduct the same study in Thailand. Then the authors will compare results obtained for Thailand with those for Japan.

Acknowledgments

Thanks to Mr. Hideo Ito, Mr. Ono, former assistant of Hokkaido Automotive Engineering College. This research was supported by a Hokkaido Automotive Engineering College encouragement subsidy.

References

[1] World Motor Vehicle Statistics vol.12, Japan Automobile Manufacturers Association, Inc., 2013.

[2]http://www.jama.or.jp/industry/industry/industry_1g1.html, Japan Automobile Manufacturers Association, Inc., 2013.

[3]http://www.jama.or.jp/industry/industry/industry_3g1.html, Japan Automobile Manufacturers Association, Inc., 2013.

[4]http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx, International Monetary Fund, World Economic Outlook Database, April 2013.

[5]http://www.stat.go.jp/data/nihon/zuhyou/n0200100.xls, Statistics Japan, 2013.

[6] http://www.airia.or.jp/number/mycar.html, Automobile Inspection & Registration Information Association.

[7] The 2006 Annual Report of Automobile Service, Japan Automobile Service Promotion Association, 2006.

[8] The 2011 Annual Report of Automobile Service, Japan Automobile Service Promotion Association, 2011.

An approach of Software Development Support in Reference to Organization Management Paradigm

Noriko Wada† Michitaka Kosaka‡

Japan Advanced Institute of Science and Technology

Abstract

We have seen great variability in the outcomes of activities to improve the capabilities of software development, and that such variability has been attributed primarily to the skills and experiences of promoters of individual support activities. In this paper, we propose an approach of software development support activity that enhance the probability of achieving expected result with reference to service approach and organization management paradigms.

Keywords, software process improvement, organization capability, service approach, organizational paradigm, SEPG

1 Introduction

In general, the abilities of software development organizations means the following two items.

1) Scientific and technical knowledge and skills of development area
2) Software engineering knowledge and skills, such as project management, capability maturity model, requirements development/management, verification and validation

In this paper, software development support activities means improving abilities of (2) software engineering knowledge and skills mainly, and activities to improve these abilities are analyzing the organization's problems from the point of view of software development processes, and incorporating solutions into development processes. These activities are generally conducted by members called SEPG (Software Engineering Process Group)[(1)] providing support to software development organizations so that together they can achieve outcomes such as solving the organization's problems or improving development abilities of the organization. In these activities, it is said that involving organization members and understanding the context of organizations are important[(2)]. However, the way of involving members and understanding organizational context haven't been discussed until now. Consequently, outcome of improving activities are greatly dependent on the SEPG individual(s) involved.

On the other hand, it is reported that software development support activities become more effective and are more certain to have desired outcome by using a service science approach[(3)] that has been investigated heavily in recent years. However, there hasn't been enough discussion about organization characteristics and how to provide support activities according to organization characteristics.

Based on such considerations, this paper proposes a methodology of identifying and delivering support services to software development organizations to improve their abilities that is suitable to the organization persona by personifying software development organization, making use of both organization characteristics which impact support activities as well as Constantine's organization paradigms[(4)].

In Chapter 2, we clarify organization characteristics which impact support activities and propose four organization personas of software development organizations in reference to Constantine's organization paradigms[(4)]. In Chapter 3, we propose the methodology to provide support service to improve abilities of software organizations that is suitable to the organization persona. In Chapter 4, the

methodology proposed in this paper is investigated by using real cases of development support in the past.

2. Software development support activities and organization persona

2.1 Significance of organization characteristics in software development support

Software development support activities are generally carried out in four steps as follows.

Step 1: Identify problems and analyze causes

Identify problems of the organization, and analyze causes. Specifically, with respect to problems such as poor quality or cost overruns, causes are identified such as lack of skills of project management and requirements management, inadequate quality assurance activities, or mismatch of development method with business model.

Step 2: Identify abilities in areas to be developed

Identify the areas in which organizations should improve to eliminate causes that are analyzed in Step 1, and set the goal of improvement.

Step 3: Develop strategies for support activities

Develop strategies to achieve the goal set for each area in Step 2. It is said to be important to consider organization culture and its strengths and weaknesses to develop strategies for carrying out activities to achieve the goal efficiently and effectively.

Step 4: Perform support activities

According to the strategy developed in Step 3, develop a concrete activity plan and execute the support activities.

In general, after having verified the result of Step 4, go back to Step 1 to improve further.

Table 1. Activities on each step of development support

Step1	
Identify problems and analyze causes	・Identify problems in the organizations ・Analyze the root causes ・Clarify the goal
Step2	
Identify abilities in areas to be developed	Identify areas in which organizations should improve and set the goal
Step3	
Develop strategies for support activities	Develop strategies for efficient and effective support activities
Step4	
Perform support activities	Develop an activity plan according to the strategy and perform support activities

To ensure that these four steps proceed, the SEPG provides support activities to the software development organization, and perform activities to improve their abilities together with members of the organization. As a result, the software development organization realizes outcomes like resolved problems or improved abilities. When progressing through these four steps, the outcomes are greatly dependent on SEPG because of the complexity of these activities at the present. On the other hand, it is reported that software development support activities become more effective and are more certain of getting desired outcomes by using a service science approach[(3)]. This means that software development support activities are services to software development organizations being provided by SEPG, adding to traditional steps of software development support activities new steps of personifying software organizations to understand the organizations' environments and real needs. It makes support activities more appropriate to characteristics of a target organization, consequently improvement activities become more efficient and effective.

Fig. 1. Definition of software development support as service

This approach is not only identifying abilities to improve but also understanding organization characteristics and providing support activities according to their characteristics. This means when treating a patient, to understand patient's characteristics and customs with the effect of a

disease makes treatments more effective.

According to this idea, it is important to understand what the organization truly demands, and to understand this fully, it is crucial to understand the organization characteristics.

2.2 Software development support and organization characteristics

In this chapter, we try to clarify organization characteristics that impact software development support activities with reference to Table 1 activities on each step of development support.

In Step 1, problems are either taken up by organization management from a medium- to long-term perspective, or identified in daily development work by members. These problems identified somewhere in the organization are taken up at the level of organization and action is taken to shift problems into issues that should be dealt with by the organization. Furthermore, the ideal situation to attain once these problems are resolved is considered at the organization level.

From these actions, organizational problem recognition process and consensus building process are executed in Step 1 to determine how problems will be identified by whom in which role, and then how these problems will be taken up as an organizational issue. In addition, consensus building is executed regarding the ideal situation for the organization, which is greatly influenced by what the organization sets a high value on.

In Step 2, the ideal situation described in Step 1 is used to identify what kinds of abilities are needed and to what degree the abilities are lacking. In these activities, organizational consensus building process and decision making process are executed to make organizational judgment.

In Step 3, organizational strategies are developed to improve abilities which were identified in Step 2. In this step, organizational decision making process and consensus building process are executed because of the importance of organizational agreement and decisions on strategies. In addition, regarding the content of the strategies, it is important to consider the organization's flexibility to change and what the organization sets a high value on.

In Step 4, strategies developed in Step 3 are substantiated to an activity plan, which is then executed. During these activities also, organizational decision making process and consensus building process are executed when organizational agreement or decision are needed.

From these, it can be seen that the following five organizational items have a great impact on how to perform support activities and on their outcomes.

- Problem recognition process
- Decision making process
- Consensus building process
- What the organization sets a high value on
- Flexibility to change

In other words, these five items are main characteristics which should be considered to perform support activities to improve abilities of the software development organization.

2.3 Personification of the software development organization

It is said that software development organizations can be divided into four groups of Closed, Random, Open, Synchronous based on the group cohesion and intrinsic flexibility[4]. In this paper, we personify software organizations by using software organization management paradigm and the organization characteristics described in 2.2.

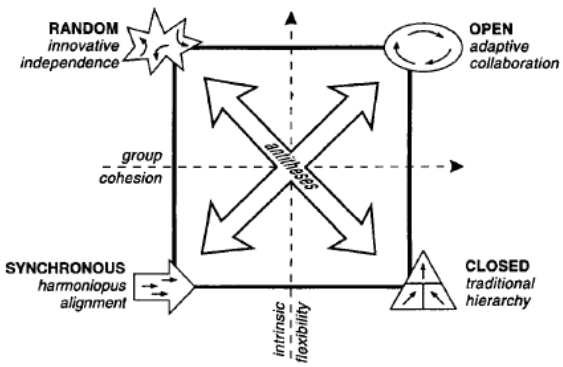


Fig2. Map showing relationships among reference organizational paradigm

(1) Closed

They have strong cohesion and weak flexibility. This type of organization is structured as a pyramid, where members tend to behave based on superior's directions. They have well-defined

roles specified for each position in the hierarchy, and information is carefully controlled. They tend to value loyalty to the organization

(2) Random

They have weak cohesion and strong flexibility. This type of organization is managed based on the independent initiative and values of individuals. They communicate with each other having respect for each other and for diversity.

(3) Open

They have strong cohesion and flexibility. This type of organization is managed by exchanging a lot of information between members, thereby building consensus.

(4) Synchronous

They have weak cohesion and flexibility. They place importance on shared and accepted common vision and discipline rather than values held by individuals. This organization is managed by members' harmony while maintaining mutual distance.

From these ideas and based on the software organization management paradigm, organization persona can be derived from the organizational characteristics described in 2.2 which impact improvement of software development abilities as shown in Table 2.

3. A proposal of a methodology to provide development support suited to organization persona

In the following sections, for each software development organization persona are shown the way to effectively and efficiently conduct activities to improve development abilities, the characteristics of development processes which are a part of the outcome of these activities, and the way of providing the required support services.

3.1 Organization persona 1: Closed type

Because Closed type persona organizations tend to place greatest value on organizational approval, for conducting support activities, it is important to obtain organization management approval at milestones and to make this approval public. It is also important to establish the approved structure to conduct support activities with well-defined roles, responsibilities and authority.

In the development processes and rules of the organization defined as a result of improvement activities, it is necessary to emphasize the timing of decisions, what should be decided, and the appropriate position of the members who make each decision, to ensure that decisions are made by individuals in accordance with their role in the organization. It is effective and efficient to make public the new development processes and rules by top management's top-down announcement.

From these points, what are expected from support services are understanding management's intention and providing sufficient and objective information to management for adequate decision making in the phase of identifying problems and developing strategies After developing strategies, it is important in conducting activities to guide the organization to ensure appropriate execution of the approval process at the right level in accordance with organizational hierarchy, by taking care to collect and provide information to decision makers in each level. This means support service providers (SEPG) are expected to behave as staff to support decision making by organization top management.

Table 2. Defining characteristics of four organizational reference paradigms

Organization persona	1: Closed	2: Random	3: Open	4: Synchronous
Problem recognition	Hierarchy	Initiative	Process	Alignment
Decision making	Formal, top-down	Informal, bottom-up	negotiated consensual	Unnegotiated, predefined,
Consensus building	Hierarchy	Initiative	Process	Alignment
Priorities	Stability, group	Variety, individual	Stability and change	Harmony, mutual identification
Intrinsic flexibility	Rigidity	Flexible	Flexible	Rigidity

3.2 Organization persona 2: Random type

Random type persona organizations tend to place greatest value on shared knowledge of essential significance rather than formal approval activities because they tend to attach a great deal of importance to individuals and diversity. Therefore the suitable approach for this type of persona is to develop and deploy solutions using a working group approach with a key member as leader who is reputed to have a great deal of knowledge about that problem area and members whose work are related to that problem area. To make public the development processes and rules, it is effective that working group members, instead of management, explain these processes to organization members including context and significance to obtain full understanding of them. In development process definitions, this type of organization tends to buy into processes that place a greater emphasis on individuals or groups such as design reviews. On the other hand, formal approval activities and reviews should be minimized, such as only at phase reviews.

Consequently, what are expected from support services are primarily identifying key members in the organization and facilitating working groups while maintaining emphasis on the key member's intension. In addition, SEPG must take care to encourage attention and communication in order to ensure that formal and informal communication paths are in place to transfer necessary information certainly from working group members to related members in the organization.

3.3 Organization persona 3: Open type

Open type persona organizations tend to have strong loyalty to the organization, are flexible, and place greatest value on consensus. Consequently, for conducting support activities, gathering opinions from every related member, discussing these, and coming to agreement are important. They prefer a working group approach in considering solutions because of a relative lack of awareness of organization hierarchy. Members of the working groups are more effective when composed of volunteers rather than by management's designation. It is important to share information about working group status within the organization, and receive organization members' opinions to be considered by the working group. When an organization decision is needed, it is effective to establish a forum such as an explanatory meeting to collect organization members' opinions regardless to organization hierarchy before making a final decision.

As for development processes and rules, it is often preferred to manage with milestone reviews done primarily by experts, rather than executing formal approval processes.

Consequently, what are expected from support services are discovering volunteers who have problem awareness, facilitating working groups, and promoting information sharing and opinion gathering/analysis within the organization about working group activities. Thus, it is important to focus on supporting the timely achievement of organizational consensus.

3.4 Organization persona 4: Synchronous type

Synchronous type persona organizations tend to place greatest value on intrinsic and fundamental matters rather than adhoc approaches because they have a shared organizational vision and perform activities with discipline based on a unified set of values. When conducting improvement activities, upon achieving a shared understanding of context, significance, and purpose, it is important to proceed while ensuring consistency with organization vision, existing discipline, and existing organizational processes. Consequently, at the start of improvement activities, attaining member's understanding of context, significance, and purpose for starting the activities is essential. In developing solutions, a working group approach is appropriate with a leader who is well trusted by management and members selected from within the organization. The status of working group activities should be openly available for reference by related members, for instance on an internal web site. Process definitions and rules would be made more effective by including process context, significance, and purpose in addition to process descriptions.

Consequently, what are expected from support

services are to support to arrange briefing session at the start-up period to make members understand the needs of improvement activities through explanation and sharing of context, significance, purpose, and goal state of the activities. When developing processes and rules, attaining organization members' understanding on the influence of the change on individuals and the context and reason for changing the processes is important. Consequently, it is essential to openly publicize the status of working group activities during process development, and to ensure that a briefing session is held at the completion of a draft version of the process documentation.

4. Analyzing proposed methodology with cases

4.1 Case 1: Closed type organization

(1) Overview

- Scheme for target organization
  About 50 members are involved to develop a business system.
- Goal and objectives
  To achieve CMMI level indicated by quality assurance group of head quarter.
- Background
  Shorten development period and quality increase are described in mid-range plan, however they didn't break down to detailed activities. According to head quarter's directions, they started improvement activities.
- The structure of activities
  Support promoting group with a few full-time members are organized just under the organization managements. They have a weekly meeting with development projects to share status of activities, and report to managements regularly.
- The strategy and a plan
  Support promoting group defined development processes and rules refer to CMMI. The contents of definitions were discussed with development projects at weekly meeting. Mini-appraisal were held regularly to understand gap with CMMI. These activities status were reported to managements regularly, and they promoted activities according to managements' indications.
- The result
  Achieve CMMI level on schedule.

(2) Analyze with a view of organization paradigm

This organization seems to be closed type because members tend to value on achieve goals at hand along managements' directions. The approach that the support group develop processes and rules according to CMMI and

Table 3. Organization paradigms and development support activities

Organization persona	1: Closed	2: Random	3: Open	4: Synchronous
Approach of improvement activities	Formal structure approach with top management approval	Working group based approach with leadership of initiatives.	Voluntary working group based approach.	Working group based approach with leadership of management.
Development processes and rules	Defining decision points and approval process explicitly	Emphasis on discretion of individuals or leaders	Defining milestone reviews with influential individual	Add background and objectives sufficiently
Key points of support activities	• Emphasize the management's intention • Steady implementation of approval process • Behave as top-management staff	• Identify initiatives • Facilitate working group activities • Take care of communication paths	• Facilitate Working group activities • Share about status of activities • Organize meetings to collect opinions from organization members	Pursue steady implementation of activities to sharing backgrounds, objectives and status of activities'

deployed to development projects with managements' approval is the approach for closed type of proposed methodology in section2. Therefore this case is providing services for closed type organizations to the closed type of organization.

Form this case, it is said that the closed type methodology we proposed is demonstrated.

4.2 Case 2: Random type organization

(1) Overview

• Scheme for target organization

A business group consisting of 3 divisions to develop embedded software. 300 people are involved. Their business climate isn't bad. The managements have successful experiences in the past and tend to have little appreciation for change. On the other hand, some leaders and members have problem consciousness strongly and took action to improve independently.

• Goal and objectives

To achieve CMMI level

• Background

Top management's intension that development ability must be improved to deal with software complexity and size are getting bigger.

• The structure of activities

The support group was organized just under the top management with some members from outside of organization and a few members form this business group.

• The strategy and a plan

The Support group promoted activities according to IDEAL model. They promoted activities according to top management's intension and reported regular basis to share the status of improvement activities among managements.

The support group defined development processes according to CMMI, got approval of managements and deployed to development projects.

• The result

Activities were continued for two years and process documents and templates that defined the outline of the development process were produced. However, as they were not referred to completely, these activities could not lead to any great successes. On the other hand, some leaders and members who have strong problem consciousness starts group working activities to define development processes for their projects. They get some kind of improvement..

Activities were stopped due to changes in top management.

(2) Analyze with a view of organization paradigm

The upper managements seemed to be closed type, but at development projects revel, they were random type with independent leaders centered. Support activities were closed type approach that organized support group under top management and emphasis reports and approval with managements. However this organization seems to be a Random type organization because of some leaders and members were independent strongly and took action to improve.

Consequently this case is that support group provide closed type service to random type organization. The reason why these activities could not lead to any great successes in spite of process documents and templates were produced, is support activities were not suitable to organization persona. On the other hand, they got success with working group activities. From this, effectiveness of Random type approach of methodology in Chapter2 were demonstrated.

4.3 Case 3: Open type organization

(1) Overview

• Scheme for target organization

About 150 members embedded software development organization.

• Goal and objectives

To create a environment that engineers develop software in a proactive manner.

• Background

Competition was intense, although the target organization's technical level was top-class. It was very difficult to continue handling tough requirements and demanding schedules. Therefore, a perception gap between management and members was generated

• The structure of activities

The support team consisted of head office members independent from development organization to analyze their problems. With

the result of analyze activities, working groups with leaders in the organization started to consider solutions and deploy them to the organization.

• The strategy and a plan

The support activities were designed to create a good atmosphere where managers endorsed member's activities to solve problems that members thought could be solved. This is because the support team assumed that the root causes of problems were due to the gap in organizational perception. Therefore, support members started by interviewing members. From the interviews, support members made a judgmental decision that the volunteers working group approach would be effective, because members were very conscious of their work and also exchange constructive opinions frequently. The support group provided facilitation of the working group and arranged regular basis meeting to report the status of working group activities from the members directly to reduce perception gap.

• The result

There were more results from each working group than were expected. Furthermore, continuing to the following step was also proposed by working group members, and they continued activities to improve. The development process was also improved, and this led to improving the capabilities of the target organization

(2) Analyze with a view of organization paradigm

This organization is evaluated as open type organization form their attitude of exchanging constructive opinions frequently. The support approach of this case was open type approach of the methodology proposed in section 2. These approach were suitable to this type of organization because of the result of this case was more than expected. From this, it is confirmed that open type approach in the proposed methodology in section 2 is suitable to open type organization.

5. Conclusion

In this paper, we proposed and evaluated the effectiveness of a methodology of performing and providing development support activities in reference to a service approach and the Constantine organization management paradigm. From these, it can be said that this methodology can externalize a portion of an expert SEPG's skills and experiences, and increase the potential for development support activities to achieve their expected result in order to deal with the problem that the outcomes of support activities for improving the abilities of software development organizations are heavily dependent on the skills and experiences of SEPG individuals.

We intend to continue verification of the methodology we proposed and conduct more intensive study to refine it.

References

[1] Mark C. Paulk, Charles V. Weber, Suzanne M. Garcia, Mary Beth Chrissis, Marilyn Bush : Key Practices of the Capability Maturity Model Version 1.1, Technical Report, CMU/SEI-93-TR-025, February 1993

[2] Bob McFeeley : IDEALSM: A User's Guide for Software Process Improvement, CMU/SEI-96-HB-001

[3] Akio Kameoka supervised : Service Science, NTS (2007) (in Japanese)

[4] Larry L. Constantine; Work Organization Paradigms for Project management and organization, Oct 1933/vol.36 No 10, Communication of ACM

[5] Noriko Wada, Michitaka Kosaka; A Methodology to Support Software Development Based on Co-creation of Service Value , ACIS12, Siem reap, Cambodia (2012)

[6] Wada, Norimatsu, Kishi; Analysis of SPI strategies and their modeling – exploring art of SPI promotion, EuroSPI2012, Austria(2012)

[7] CMMI Product Team: CMMI for Development, Version 1.3, Technical Report, CMU/SEI-2010-TR-033, November 2010

[8] Koichi Nagachi, Hirotake Goto: A case of SPI community approach - A realization of cross-organizational innovative change -, SPI-Japan08, Kobe, Japan, (2008)

[9] So Norimatsu: Solution focus and model focus –Searching suitable process improvement mood for the organization–, SEPG Japan 2004, Tokyo, Japan, (2004)

Building Knowledge of Customer Distinction for Safe Transaction in the Mail Order Industry

Hiroaki Azuma† Masakazu Takahashi‡ Kenichi Yoshida* Kazuhiko Tsuda*

†Hazs Corporation
‡ Graduate School of Innovation Technology Management, Yamaguchi University
* Graduate School of Business Sciences, University of Tsukuba
hirokiazuma2002@hazs.biz
maskazu@yamaguchi-u.ac.jp
yoshida@gssm.otsuka.tsukuba.ac.jp
tsuda@gssm.otsuka.tsukuba.ac.jp

Abstract

This paper presents investigating the customer characteristics of mail order industry, especially the bad debt customers. These kinds of investigations have not made intensively, such as the delivered address or name so far and conventional method for predicting such information depends on the employee's working experiences. For these backgrounds, we observed the bad debt list gathered from a mail order company and analyzed. From the results of the research, we characterize the potential customers with the natural language process. Intensive research revealed that the characteristics of customers who ordered from the vicious address. This result will make use for the revenue expansion with the improvement of the bad debts collections.

Keywords: Mail Order, Customer Analysis, Bad Debts, Risk Management, Resilience, Natural Language Process, Service Science and Management Engineering

1 Introduction

Mail order industry is one of the promising methods of sales expansion even such as deflation condition overlap with long-term slump in Japanese retail business that has characteristics of delivering the items to the customers' hand.

A survey from the Japan Direct Marketing Association says the percentage of credit losses of a mail order company is estimated about 0.5% of net sales. This number is indicated after the collection activities, such as delivering notice of reminders or phone calls. From the results of the vicious transactions, we found the address or the name of such transactions were made hard to distinguish with the conventional information processing methods.

They well understand the function of information retrieval for the distinctions and make it hard to extract the vicious transactions so far.

In this paper, we carry out data analysis of the customer distinction as the decision support knowledge in the mail order industry with natural language process and propose the method for the customer distinction for the fraud transactions.

The rest of the paper is organized as follows: Section 2 discusses the backgrounds of the research and related work; Section 3 briefly summarize the gathered data on the target mail order company; Section 4 describes analytics of the data and presents analytical results; and Section 5 gives some concluding remarks and future works.

2 Backgrounds and Related Work

While the domestic retail business is in the long-term slump, the mail order industry is continuing the sales expansion. One of the conclusive factors of the sales expansion is the reversionary payment method. Meanwhile, it increases in number such as fraud transactions from vicious customers in recent year.

Therefore, it is important to realize safe dealings in the mail order industry, that provides redistribution to the customers and reduce sunken security costs. Moreover, to increase the transactional number of the mail order that is achieving economic growth in Japan, let the domestic economic demand stimulate.

According to the mail order industry sales survey, it was estimated amount to 5,090 billion yen in total of the 2011 fiscal year which increased in 9.0% raise compared with previous year [1]. One of the factors for sales expansion among the mail order industry is the diversification of payment methods. Not only the reversionary system that is settles accounts after the order item arrives with which improving customers' conveniences also appeal the transactional safety from the view point of the customers which expect the large increase both of the sales and the profits for the mail order company.

On the other hand, there are no credit criteria at the time of start trading. Therefore, there are many

transaction of scam or fraud by vicious customers making use of the reversionary systems.

As for the related work on the mail order industry, it separates into the activity of before order and of after order received. The related work of the before order activities were focused on the elaboration of the order received in consideration of the time lag from customers such as the demand predictions [3]-[6].

On the other hand, at the phase for the after order received, most of the researches have been made for customer analysis with the purchase history. There are lots of customer analyses with the data-mining methods [7], [8]. For example, the order history has been used for such as trend analyses of customers and sales promotion strategies as the retailers' decision support tools.

As long as the reversionary system is adopted in mail order industry, it is important to collect the payment from the customers as soon as possible, whereas the related work on the collecting the payment was less studied than that of customer analysis, especially collecting the bad debts. Concerning to the research on the bad debts, most of them have been done about the bankruptcy prediction of the corporation [9]-[11]. These are used for the preliminary screening to hedge risks with the machine learning methods [12]-[15].

Among the machine learning methods, the random forest is one of the promising methods for such an insufficient data prediction such as authorship identifications, software production workload estimations, promising customers' classifications, and credit risk evaluations [16]-[22]. Takahashi et.al made customer analysis with random forest to extract regionality [23]. Meanwhile, there is little research in the loan management so far.

Therefore, under the present circumstances, the worker who applied and became skillful about the contents performs identification of multiple accounts under the same name as a single entity visually, and personal correspondence of deducing an arrears customer is performed. The conventional method for detecting the transaction with doubt is made by the rank-and-file employees with their working experiences, so far.

3 Data Summary

There are lots of transactions that take reversionary payment systems in the mail order industry sector of the health food and the cosmetics. We make data analysis of the target sector based on the order data and payment data. We make use of the transactional data from the mail order company handles the basic skin-care.

This data is composed of about 300 thousand transactions in reversionary payment method amount to 1,800 million Japanese yen in sales. This company has around 3,000 million Japanese yen annual sale in 2012. Since this company takes the policy for the first time purchase of the limitation to use with cash on delivery or credit card, they have good score of the debt balance compared with another company in the mail order industry.

From the transactional data of the target company, we take the categories as follows; 1) Customer number, 2) Order number, 3) Customer name, 4) Customer name in kana, 5) Zip code, 6) Address 7) Phone number, 8) Order date, 9) Payment date, 10) Purchased item names, and 11) Prices. We make intensive comparison both of the ordinary transactions and the accrued one.

From Table 1. to Table 4. indicate the results of the analysis, respectively. From the results, the ambiguous personal information such as names and addresses for escaping specifying the individual have frequently observed.

Table 1. indicates the examples of false names. There are frequently observed the address expressions such as used the maiden name, given name of sister or children. And the combination of name of children and the c/o address is used for the fraud transaction to order repeatedly without the machinery detection. In the center of Tokyo, there are frequently observed that the door plate is not displayed at the mail box of apartment, which is make ill use of the fraud delivery since the assignment of Logistics Company could finish delivering the address of the destination.

Furthermore, making use of completely different name to the destination or the combinations of letter with Hiragana, Katakana and Alphabetic character are frequently observed among the list.

Table 1. Use c/o as the Delivery Address

Name	Date of Order	Zip+Address Code	Address
Keiko Saitou	1/21	30300411541	Takahashi*SAMAKATA*, 1541, Hei, Tanaka Town, Satou City
Matsumoto Kazuko	1/21	30300411541	Takahashi*SAMAKATA,* 1541, Hei, Tanaka Town, Satou City
Keiko Saitou	2/20	30300411541	Takahashi*SAMAKATA*, 1541, Hei, Tanaka Town, Satou City
Matsumoto Kazuko	2/20	30300411541	Takahashi *SAMAKATA*, 1541, Hei, Tanaka Town, Satou City

Keiko Saitou	3/21	30300411541	Takahashi*SAMAKATA*, 1541, Hei, Tanaka Town, Satou City

Table 2. Use Different Building Name as the Delivery Address

Name	Date of Order	Zip+Address Code	Address
Mari Wada	3/14	9550842213182	*CORPO Yoshida* 2, 2-13-18, Tori, Itou City
Masami Wada	3/26	9550842213182	*Yoshida CORPO* 2, 2-13-18, Tori, Itou City

Table 3. Use Different Type Set and Deleting Building Name as the Delivery Address

Name	Date of Order	Zip+Address Code	Address
Mai Matsuda	2/22	886085369102	*CITYCOURT F102, 69,* Aobanishi, Kadowaki Aza , Oonishi City
Yumi Matsuda	7/20	886085369102	*F102, 69,* Aobanishi, kadowaki, Oonishi City

Table 4. Use Address Features as the Delivery Address

Name	Date of Order	Zip+Address Code	Address
Natsumi Nakamura	4/16	260156127	127, Michishita *No* machi, Suzuki City
Tomo Kobayashi	4/18	260156127	127, Michishimo town, Suzuki City
Fumi Katou	4/18	260156127	127, Shimo Town, MIchi Aza, Suzuki City

Table 2. indicates the example of different building name as the delivery address. Among the list, Mixed up the composition of apartment name, Transform the apartment name into another character set such as Hiragana from Kanji, and Synonymous Japanese expressions that are making ill use of English pronunciation are also observed.

Table 3. indicates the use different type set and deleting building name as the delivery address. Among them, no indication of the apartment name, Name of the apartment such as Mansion is simply abbreviated to MS, and abbreviation room number is observed.

Table 4. indicates the use Address Features as the Delivery Address. Some characteristic address expressions are observed as follows; Attach the excessive information in the address such as 1) Unique Kanji; "Aza"or "No", 2) east building or west building that do not exist, 3) Former address notation, 4) Replace Kanji into the numerical value seen in Hokkaido or Kyoto, 5) Unique address notation of "Up" or "Down" seen in Kyoto, 6) Delete hyphens, 7) Change hyphens into Kanji, 8) Change locations 9) Change Room expressions like Room 101 into 1F right, 10) Cover them up like apartments.

From the result of analysis, about 1 million Japanese yen out of 67 million yen, 1.49% of the Sales are extracted from the proposed method.

4 Data Analysis

In order to make fraud transactions with success, they usually use intentionally the fictitious names and the synonymous addresses to avoid distinction. Therefore, we make use of the bad debt customer list after the customer identity screening with the numerical information. As the bad debt customer definition, we focus the target customers as following definitions; the address is same but the customer names are different and their payments are not competed yet.

The analytical procedure is described as follows;

Step1 Analytical Data Preparations

1-a) Merge the customer address and zip code for the new address information.

1-b) Sort the addresses to capture the same address.

1-c) The random masking program manipulation was processed to hide the individual specification in consideration of the private information protection Law.

Step2 Data Analysis
Each address feature was extracted from the list with the following procedures.

Apartment name difference extraction
2-a) Make the group both the address number and address is different.

2-b) Extract the data of apartment name is different from the group.

Name differences
2-c) Extract the Hiragana name after sorting the name list.

c/o, "*Aza*" and First character of the beginning is Hiragana, Katakana, and English

2-d) Extract the target characters after sorting the name list.

2-e) Aggregated the above extracted data

Summarized data is described in Table 5. This table indicates the comparison of the paid customer and outstanding, respectively.

There is definitive difference both between the paid customers and the others in the address description. In case of no statement for the apartment name, the ratio of the outstanding increases about 1.74 times compared with the ordinary transactions, as a result.

Table 6. and Table 7. indicate the comparison results of the paid customers and the accrued customers in names. Table 6. shows the comparison results of the name expression. Generally, the recipient names written in Hiragana or Katakana have high frequency of delinquency for the payment, whereas this trend is not seen from the analysis. As the example of "Hiroki Azuma"'s notation, "HIROKI AZUMA", "Hiroki AZUMA", "HIROKI Azuma", "hiroki azuma", "Azuma Hiroki" can be considered. However, the method of making Hiragana and Katakana intermingled for recipient addresses have possibility of intentionally from the arrears customers.

Table 5. Difference between the Paid Customers and the Accrued Customers in Address

	Paid (a)	Accrued (b)	(b) / (a)
Name of the apartment	0.1146%	0.2000%	1.7452
Name in Hiragana	0.0244%	0.0405%	1.6598
c/o	0.6302%	0.9082%	1.4411
Hiragana, Katakana, English	0.1661%	0.2226%	1.3402

Table 6. Comparison for the Paid Customers and the Accrued Customers in Name

	Paid (a)	Accrued (b)	(b) / (a)
Katakana	0.4759%	0.6765%	1.4214
Hiragana	0.0644%	0.3031%	4.7055
Katakana＋Hiragana	0.1769%	0.8081%	4.5682
Katakana＋Kanji	0.0617%	0.0476%	0.7718
Hiragana＋Kanji	0.0118%	0.0054%	0.4565

Table 7. Number of Characters in Name

Number	Paid (a)	Accrued (b)	(b) / (a)
1	0.0006%	0.0006%	0.9767
2	0.2558%	0.4991%	1.9508
3	12.3749%	14.8433%	1.1995
4	123.9998%	122.1695%	0.9852
5	60.0468%	58.5415%	0.9749
6	2.7607%	3.6786%	1.3325
7	0.4285%	0.1451%	0.3385
8	0.0889%	0.1108%	1.2471
9	0.0268%	0.0099%	0.3686
10	0.0059%	0.0017%	0.2930
11	0.0042%	0.0000%	0.0000
12	0.0036%	0.0000%	0.0000
14	0.0006%	0.0000%	0.0000
15	0.0018%	0.0000%	0.0000
16	0.0012%	0.0000%	0.0000

Table 8. Characteristics of the Address

Elements contained in the address	Paid (a)	Accrued (b)	(b) / (a)
include "*Kata*" or "*Samakata*"	0.3256%	2.5057%	7.6968
include "*Hidari*"	0.2196%	0.3177%	1.4464
include "*Mig*"	0.2550%	0.1447%	0.5675
include "*Sayu*"	0.4746%	0.4624%	0.9742
complete "*Shitsu*"	0.0038%	0.0172%	4.5889
contains "*KannrininShitsu*"	0.0009%	0.0077%	8.8379
include "*Nai*"	1.3573%	2.7127%	1.9985
include "F"	0.8716%	2.8353%	3.2531
include "*Kai*"	0.4494%	1.3177%	2.9324
include " ."	0.2300%	0.5789%	2.5167
include "--"	0.0020%	0.0057%	2.8408
complete except number or the characters	2.0171%	11.4298%	5.6664
include "*Aza*"	3.3928%	4.0214%	1.1853
include "*Aza・No*"	4.3952%	5.8088%	1.3216
include "*Chuou*"	2.0321%	2.7703%	1.3633

Table 9. Address Examples

1	Taro Yamada Samakata, 1-4-3, ●●Town
2	Corpo Midori 2F Migi, 2-8-2, ▲▲
3	Asahisou A-2F Hidarigawa, 2-6-14, ◆◆Town
4	☆☆Palace Ⅱ202Shitu, 4-15-13, ★★
5	Green heights Kanrininshitsu, 5-17-22, ○○Town
6	□□Shoji Nai, 3-18-1, ■■
7	▽▽Heights 1F2, 1-6-1, △△
8	◎◎Apartment 1Kai, 2-15, ◎◎Town
9	1-206, 4-6-26. , ***
10	**Apartment 1Kai, 1-5--26, **Town

Table 7. indicates the comparison results of the number of characters in names. As for the count method of the character, " Hiroki Azuma " counts three characters and " HIROKI AZUMA " counts six characters, respectively.

The total number of characters in name is compared. The average number of characters in family name up to the 24,706th place is 2.09 characters from the national raking database of the Japanese family name [23].

From the analysis, it turned out that the rate of arrears becomes high in case of the number of the character in name less than the average.

Table 8. and Table 9. indicate the comparison results of the paid customers and the accrued customers in address.

Table 8. indicates the characteristics of the addresses. Table 9. indicates the example of the address. Characteristic address expressions are found from the analysis such as include "*Kata*" or "*Samakata*" , "*Hidari*" , "*Migi*" , "*Sayu*" , "*Shitsu*" , "*KanrininShitsu*" , "*Nai*", "F" , "*Kai*"" ,".","--",and complete except number or the characters in the recipients address.

5 Concluding Remarks

In this paper, the extraction of the characteristics in name and address with the comparison of the white and black data is performed. As a result, we are succeeded in extraction of the characteristics in name and address.

As for the name expression, it turned out that the rate of arrears becomes high in case of the number of the character in name less than the average. and as for the address expression, Characteristic address expressions are found from the alasisys such as include "*Kata*" or "*Samakata*" , "*Migi*" , "*Hidari*" , "*Sayu*" , "*Shitsu*" , "*KanrininShitsu*" , "*Nai*" , "F" , "*Kai*"" ,".","--",and complete except number or the characters in the recipients address.

Our future work includes 1) analysis regarding analysis regarding customer attributes and purchased categories, 2) generate new transactional patterns with meta-heuristic algorithms, and 3) mount to detection system as decision support system for the mail order industry.

These works will require practical experiments and further survey studies.

Acknowledgment

We wish to express our gratitude of the cooperation from the mail order corporation to our analysis.

References

[1] The Japan Direct Marketing Association: 19th Annual National Mail Order Survey Report, 2012(Japanese)

[2] Simester, D. I., Sun, P. and Tsitsiklis J. N.: Dynamic Catalog Mailing Policies, Management Science, 52(5), 683–696,2006

[3] Kimijima M.: A Study on Measuring Input-Output Process on Order-Getting Costs for Direct Marketing, Yokohama National University Departmental Bulletin Paper, 16(1), 21-39,2010 (Japanese)

[4] Conlin M., O'Donoghue T. and Vogelsang T. J.: Projec-tion Bias in Catalog Orders, American Economic Review, 9(4), 1217-1249, 2007

[5] Matsuda Y., and Ebihara J.: Forecasting Model in the Mail-Order Industry, UNISYS Technology Review, 71, 52-68, Nov., 2001 (Japanese)

[6] Motoda H., and Washio T.: Perspective of Data Mining, System/Control/Information, the Institute of Systems, Control and Information Engineers, 46(4), 169-176,2002 (Japanese)

[7] Ishigaki T., Motomura Y., and Chan H.: Consumer Behavior Modeling Based on Large Scale Data and Cognitive Structures, IEICE Technical Report, NC2008-157, 108(480), 319-324, 2009 (Japanese)

[8] Yamashita S., Tsuruga T., and Kawaguchi N.: Consideration and Comparison about a Valuation Method of a Credit Risk Model, Discussion Paper Series (11), Financial Research Center, Financial Service Agency, 2003 (Japanese)

[9] Yano J., Ikai M., Nakagawa K., Takahashi S., and Namatame T.: A Study on Credit Card Users' Default Prediction, Journal of Operations Research of Japan, 51(2), 104-110, 2006 (Japanese)

[10] Sunayama W., and Yada K.: Modeling and

Analysis of Persuading Process using Conversation Logs, Proc. of the 20th Annual Conference of the Japanese Society for Artificial Intelligence, 3C3-3, 2006 (Japanese)
[11] Tanabe K., Kurita T., and Nishida K.: Prediction of Corporate Credit Ratings by Support Vector Machine, Journal of Japan Society for Management Information, 20(1), 23-38, 2011 (Japanese)
[12] Min J. H., and Lee Y.C.: Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters, Expert Systems with Applications, 28(4), 603-614, 2005
[13] Abe N., Melville P., Pendus C., Reddy C.K., Jensen D.L., Thomas V.P., Bennett J.J., Anderson G.F., Cooley B.R., Kowalczyk M., Domick M., and Gardinier T.: Optimizing Debt Collections Using Constrained Rein-forcement Learning, Proc. of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 75-84, 2010
[14] Hashimoto M., and Yoshida K.: A Study on the Evaluation Technique of the Claim Assessment Model in a Retail Financial Sector, Proc. of The Annual Conference of The Japan Society for Management Information, 63-66, 2004 (Japanese)
[15] Breiman, L.: Random Forests, Machine Learning, 45(1), 5–32, 2001
[16] JIN M., and Murakami M.: Authorship Identification using Random Forests, Proc. of the Institute of Statistical Mathematics, 55(2), 255-268, 2008 (Japanese)
[17] Jin M.: Estimation of When the Works were Written: With the Works of Ryunosuke Akutagawa as Examples, Behaviormetrics, 36(2), 89-103, 2009 (Japanese)
[18] Kobayashi Y., Tanaka S., and Tomiura Y.: Classification and Assessment of English Scientific Papers Using Random Forests, IPSJ SIG Technical Reports, 2011-CH-90(6), 1-8, 2011, (Japanese)
[19] Konishi F., Uchida S., Toda K., and Monden A.: An Approach of Software Development Effort Estimation by Random Forests, Proc. of the 2011 IEICE General Conference, Information/System(1), 17, 2011 (Japanese)
[20] Ohashi K., Toyoda H., and Kubo S.: A Study on Classification and Prediction of the Potential Customers, Journal of Operations Research of Japan, 56(2), 71-76, 2011 (Japanese)
[21] Umezawa Y., and Mori H.: Credit Risk Evaluation of Power Market Players with Random Forest, Trans. of the IEEJ, Power and Energy Society, 128(1), 165-172, 2008 (Japanese)
[22] Takahashi M., Azuma H., Ikeda M., and Tsuda K.: Building Knowledge for Characterization of the Bad Debt Customers in the Mail Order Industry with Random Forest, Advances in Knowledge-Based and Intelligent Information and Engineering Systems, Frontiers in Artificial Intelligence and Applications, 243, 867-847, IOS Press, 2012
[23] Japanese sure name ranking database, Shizuoka Univ. available at http://www.ipc.shizuoka.ac.jp/~jjksiro/kensaku.html (accessed May 1st, 2013)

ACIS 2013

Asian Conference on Information Systems 2013

Special Sessions on Computational Medical Image Processing

Active Contours (Snakes) and their Applications to Medical Images. A Mini-Lecture

Stanislav S. Makhanov
School of Information and Computer Technology, Sirindhorn International Institute of Technology, Thammasat University, Thailand
makhanov@siit.tu.ac.th

Abstract

The paper offers a short survey of the snake-based segmentation methods employed by computerized medical image processing along with relevant pre-processing such as speckle noise filters, region growing and clustering methods.

1. Introduction and Literature Survey

The active contour evolves to minimize the image energy which includes the internal energy of the contour and the external energy of the image. The minimization is governed by the corresponding Euler equations involving the internal and external force. The internal force maintains the contour smoothness, while the external force attracts the contour to the desired boundaries. The deformation stops when the snake achieves a minimum of the energy. The main drawback of the method is that the noise and small objects may attract the snake to a local energy minimum, which does not correspond to the actual boundary. Therefore, to reach the desired boundary, the initial contour should be initialized close to the object. The most important component of the snake based segmentation is the external force which pushes the snake towards the object. This force is usually derived from the gradient of the image gray level. Therefore, in order to enhance the effect of the external force the gradient created nearby the boundary must be extended so that the snake "feels" the object even if it is initialized far from it. A popular solution is the balloon snake where such a force is generated artificially. This works for both contracting and growing snakes (balloons or artificially inflated contours) proposed in [11]. The distance snakes [12] exploit a similar idea. The external force is constructed as the negative of the external energy gradient, which is the distance from each point to its closest edge points in the image. Consequently, the initial contour can be located far away from the desired boundary if there are no spurious edges along the way. Furthermore, many variations of this idea such as the "stop and go" snakes [13], multi -direction snakes [14], gravitation force snakes [15], watershed-balloon snakes[16], balloon snakes with nonlinear filtering [17] have been introduced and analyzed. The image force is modified or altered to increase the capture range and decrease the sensitivity to the possible noise, shadows and (in case of the medical images) obstructing structures and tissues. Another group of methods is based on minimization of the energy subject to a certain conditions improving the convergence and accuracy. Sectored snakes [18] deform the contour subject to constraints derived from an priori knowledge of the object shape, extracted from the training set of images. Fourier type descriptors have been used in [19] to evolve the curve to a prescribed shape defined by a template. The prior information is introduced through a set of invariants (translation, rotation and scaling) computed using the Fourier Transform.

Furthermore, the force includes not only the edge based features but region based features as well. For instance, the homogeneity of the enclosed region[21]. In [22] region-based image features are combined with the edge-based features incorporated in the external forces. Starting from multiple seeds, [22] performs segmentation of the entire image by iterative boundary deformation and region merging iteratively. The so-called T-snakes proposed in [23] and their improvements such as the dual T-snakes [24] based on iterative re-parameterization of the original contour are able to make the use of the self loops. However, the approach allows only "rigid" deformations limited by the superimposed "simplicial grid". An intrinsic internal force that does not depend on contour parameterization based on regularized contour curvature profile has been proposed in [25-26]. A competing approach called the level set method [27] is based on the ideas proposed by Osher and Sethian [28] to use a model of

propagating liquid interfaces with curvature-dependent speeds. The level set method combined with the contour energy minimization resulted in a variety of the so-called geodesic deformable models [29-32]. However, the level set method makes it difficult to impose arbitrary geometric or topological constraints on the evolving contour via the higher dimensional hyper surface. Besides, the level set models may generate shapes having inconsistent topology with respect to the actual object, when applied to noisy images characterized by large boundary gaps [33]. Li et al. [34], in reference to the problem of topological adaptation and computational complexity write “in light of the . . . inherent weaknesses of geometric active contour models, it is worthwhile to seek solutions within the parametric model realm". Rochery et al. propose a parametric model for higher-order active contours, in particular, quadratic snakes, for extraction of linear structures like roads [35]. The idea is to use a quadratic formulation of the contour's geometric energy to encourage anti-parallel tangents on opposite sides of a road and parallel tangents along the same side of a road. These priors increase the final contour's robustness to partial occlusions, decrease the likelihood of false detections in regions not shaped like roads, and help to prevent self-looping.

Further improvements lie along the lines of processing the underlying vector field rather then modifying the snake model itself. A number of popular codes are based on a gradient vector flow (GVF) method proposed by Xu and Prince [36],[37]. A “raw” gradient vector field derived from the image edges is replaced by a vector field which minimizes a certain variational functional. The functional is designed to extend the large gradients far from the boundary, smooth the noise and speckles while keeping gradients attached to strong edges. The corresponding Euler equations represented by linear elliptic PDEs are solved numerically. The PDEs are interpreted as the steady state representation of a non-stationary diffusion process with constant diffusion coefficients. The generalized gradient vector flow field (GGVF) [38] extends the GVF method by introducing non-uniform diffusion. The GGVF is defined as a steady state solution of a system of parabolic equations with the elliptic terms and the source term similar to the GVF model. However, the GGVF employs space-dependent diffusion which provides better segmentation accuracy and a larger capture range. Some variations of these ideas are the multidirectional GGVF based on a special algorithm to evaluate the gradients and the diffusion coefficients [39] and the non-linear diffusion method presented in [40]. Numerous research papers apply the GGVF based active contours to medical images. The examples are multi-direction snakes: skin cancer images [14], topology-adaptive snakes: MR brain images and CT scans[23], gravitational force snakes: a variety of medical and non-medical images [15], narrow-band snakes: MRI and CT scan images of lungs [41], distance snake, GVF snake, balloon snake , “area and length” snakes, geodesic snakes, constrained snakes and level set method: MRI, CT and US images of brain, liver, kidney [32], region-competition snakes (originally [22]): CT scan slices of arteries [44], sectored snakes : abdominal CT scans[5], parametric snakes: US of breast masses [45], 3D-snakes: US breast cancer images[45],[46], GVF snakes with edge map pre-processing: US of the kidney tumors [9], GVF snakes combined with the region growing and the median filter :US breast tumors [10], sketch-snakes: chest X-Ray images [46], combination of snakes and active shape models: US of the human heart [47], the so-called early-vision and the discrete-snake model: a variety of the US images [48], multi-resolution snake: echographic and echobrachial images[49], GGVF snakes combined with a continuous force field analysis: breast tumor US images [50] .

The success of such segmentations critically depends on preprocessing. As a matter of fact, it is often more important than selecting and tuning the active contour method. The Gaussian, mean and median filters [51], Gabor filters [48] and speckle noise filters [51]-[53] are among the most popular for the US imagery. A fusion of the median and the Wiener filter combined with a contrast adjustment technique. and the Frost speckle filter [53] is applied in[55] . A tree-structured nonlinear filter and special types of wavelet transforms has been proposed in [56] for transrectal ultrasound. The active contour is used by the segmentation module. A comparison between the segmented image with and without the preprocessing shows that this module produces a great improvement in the accuracy of the boundary detection. A combination of filtering, edge map and initialization by a human operator [10] employs an iterative truncated

median filter to reduce the speckle noise. In [57] the speckle noise is suppressed by anisotropic diffusion filter[58] and a stick filter [59].

More often than not, the preprocessing is not a single step but a well arranged sequence of operations including (but not limited to) filtering, morphological transformations and edge detection procedures specific for the particular type of medical imagery. A combination of region growing and median filtering is proposed for the GGVF snakes in [10]. A single run of the anisotropic diffusion filter (non-linear filter[60]) is proposed for the multi direction snake [14]. Non-linear filters applied to the GVF vector field (rather than to the original image) are discussed in [17]. A curvature diffusion filter [61] is applied with the level set method for sonographic images of breast. A human operator initializes the snake close enough to the actual boundary[62]. A similar interactive medical image segmentation (sketch snakes) is introduced in [46]. Some preprocessing operations such as the edge detection can be also controlled interactively by a human operator. A preprocessing for density-based segmentation [63] employs histogram adjustment, noise reduction using an iterative dilation [64] and a median filtering to suppress the spike-like noise. Multi scale approach combined with the active contours shows much promise as applied general segmentation [65] as well as to segmentation of the boundary of the endocardial contour in ultrasound images[66]. The idea is that a rough initial contour can be extracted at the highest level of the multiresolution structure. This crude contour is then used for the initialization of the optimization process at the coarsest level. The multiresolution techniques can be combined with a variety other preprocessing methods such as the phase portrait analysis [67] which applies directly to the vector field. In summary, GGVF endowed with an appropriate sequence of noise elimination steps remains one of the most popular choices for segmentation of the US images. The method removes gradients produced by the speckles and the tissue-related edges and allows to hand-tune coefficients and the force terms in the corresponding Euler equations. Furthermore, the parameters of the snake can be adjusted depending on the anticipated shape and size of the tumor.

However, when the noise related gradients are comparable with the boundary gradients, the diffusion smoothes the false and the true contour points equally. The smoothing after a large number of time steps (iterations) may change the direction of vectors pointing (correctly) towards each other at the true boundary. Besides, even though GGVF produces smoother vector fields it may lead to undesirable effects nearby concave boundaries such as the attracting or repelling stars. Moreover, the GGVF does not include the directions of the vectors but only their magnitude. Consequently, a weak boundary which generates a sequence antiparallel vectors having a small magnitude might be destroyed by the diffusion. Moreover, the boundary may appear only in a certain range of the gray levels of the edge map. Therefore, the GGVF must adapt to the intensity of the edge map and the direction of the vector flow as well. In this paper the adaptive diffusion coefficients are evaluated on the gray level edge map by generating clusters corresponding to the background, the shadows/noise and the boundary. The centers of the clusters are used to generate an interpolating function to find the diffusion coefficient for an arbitrary gray level. These techniques can be applied only once at the preprocessing step or during the entire course of the GGVF iterations. The second component is the adaptation to the local directions of the vector field. If the local configuration includes anti parallel vectors, the corresponding diffusion coefficient gets decreased and the “stopping term” increased. The techniques generalizes the Continuous Orientation Field Flow Analysis [50]. The previous version was based on the incremental improvement of the underlying edge map whereas the new version employs a straightforward correction of the coefficients. Unfortunately, GGVF often converges to an “over smoothed” solution. As a matter of fact, without a careful choice of the parameters many US images will be partly or even entirely smoothed by the GGVF diffusion so that only a partial boundary or no boundary at all appears on the final vector field. In this case the GGVF iterations should be interrupted even though there is still no formal numerical convergence. However, it is hard to find a theoretically justified stopping criteria. Besides the resulting snakes can generate spike-like segments when the image noise is positioned in the neighborhood of the true boundary. Therefore, we present two important correction procedures. First of all, we

modify the generic GGVF method in such a way that it admits internal boundaries where the vector flow is temporarily or permanently fixed ("frozen boundaries"). The GGVF diffusion equation is then solved in the rectangular domain with the internal boundaries characterized by the Dirichlet boundary conditions.

This modification improves convergence to meaningful solutions and circumvents the convergence to the over smoothed vector fields. Our second correction employs the Continuous Orientation Field Flow Analysis to remove the false boundaries. When the snake is close to convergence we run the COFFA window which scans the contour at every iteration. Short segments with a low COFFA-score are classified as false boundaries. At these segments the algorithm moves the snake away until it attaches itself to the true boundary. The correction step works particularly good to eliminate spike-like parts of the snake when it accidentally attaches itself to small groups of noisy pixels in the neighborhood of the true boundary. In this case the spike part of the snake is removed and the loose ends get reconnected. Finally, our efficient preprocessing method based on hierarchical fuzzy-C mean based region growing techniques provides up to 20-30% increase in the accuracy of the final segmentation. Our numerical experiments on numerous US images show a better accuracy and less sensitivity to the initial position of the contour. The proposed approach leads to a better convergence and better segmentation accuracy with the reference to the conventional GGVF snakes. The adaptive GGVF-COFFA method also performs well combined with other popular active contours such as the distance snakes and the boundary snakes. The algorithms have been tested on a set of ground truth images hand-drawn by leading radiologists with Queen Sirikit Center for Breast Cancer of Bangkok. The error is evaluated in terms of the maximum and the average Hausdorff distance(contour based measure) as well as in terms of the percentage of the true positives(pixel based measure).

2. Active Contours (Snakes)

An active contour or snake parametrically defined as $X(s)=(x(s),y(s)), s\in[0,1]$ is a curve which evolves inside the image domain so that it attaches itself to the desired object. The evolution of the snake [1] is governed by

$$a\frac{d^2X}{ds^2}+b\frac{d^4X}{ds^4}+\nabla E_{ext}=0. \qquad (1)$$

The first (tension) term establishes an equidistribution of points along the resulting curve whereas the second (stiffness) term ensures against large curvatures. The coefficents a and b define the relative importance of the tension and stiffness, whereas E_{ext} is the external (image) force which pulls the snake towards the desired object boundaries. Formally, is a linear ODE for which a closed form solution can be easily found. However, $E_{ext}=E_{ext}(X)$. Therefore, the external force depends on the solution. Therefore, only a numerical solution is possible. The simplest choice of E_{ext} is a magnitude of the gradient of the gray level subjected to a Gaussian filtering. Conventional smoothers such as the Laplacian applied to the gradient fields may remove some noise, but they also may destroy the true boundary. In particular, ultrasound (US) images of the breast tumors could be very noisy with speckles and tissue related textures obstructing the actual tumor.

3. Generalized Gradient Vector Flow Snakes

The gradient vector flow (GVF) was proposed by Prince and Xu [36-37]. The technique replaces a "raw" gradient vector field derived from the image edges by a vector field which minimizes a certain functional which extends the large gradients far from the boundary and smoothes the gradients caused by noise. The GVF is a vector-function $V\equiv\{v_x,v_y\}$ which is a solution of a linear elliptic equation given by

$$\mu\nabla^2V-(V-\nabla f)|\nabla f|^2=0,$$

where μ is the "diffusion" coefficient, ∇f is the gradient field derived from the edge map of the image. For instance, $f=|\nabla I|$, where I is the gray level of the edge map. A variety of edge detectors such as the Canny or Sobel detectors can be applied to evaluate $|\nabla I|$.

The above equation is solved by treating V as a function of a pseudo-time and by performing numerical iterations with regard to the pseudo-time. These iterations can be interpreted as solving numerically a heat equation given by

$$\frac{\partial V}{\partial t}=\mu\nabla^2V-(V-\nabla f)|\nabla f|^2.$$

However, the uniform diffusion often produces excessive smoothing. Therefore, Xu and Prince [38] extended the GVF technique by introducing spatially varying coefficients to decrease the smoothing effect, namely,

$$\frac{\partial V}{\partial t} - g\left(\left|\nabla f\right|\right)\nabla^2 V - h\left(\left|\nabla f\right|\right)(\nabla f - V) = 0 \qquad (2)$$

The improved version of the GVF is called the generalized gradient vector flow (GGVF). The weighting functions g and h depend on the gradient of the edge map so that in the proximity of large gradients g gets smaller whereas h becomes larger. In [39] the following diffusion and "attraction" functions have been proposed

$$g\left(\left|\nabla f\right|\right) = e^{-\left(\left|\nabla f\right|/K\right)},\ h = 1 - g\ , \qquad (3)$$

where K is a calibration parameter. Finally, $E_{ext} = V(x, y)$ completes the formulation of the generalized gradient flow snakes.

4. Orientation Force Field Analysis

The boundary of the object the vector field consists of antiparallel vectors. In this case the diffusion coefficient must be small and the "stopping coefficient" in the GGVF equations must be large. As opposed to that, parallel vectors corresponding to a background should entail large diffusion, so that small noise is smoothed and the large gradients along the boundary propagate through this area. The Continuous Orientation Force Field Analysis (COFFA) measures a deviation from an ideal anti-parallel position. The measure is then used to generate a score vector. In order to capture the direction along which the vectors are aligned the most, the algorithm employs a rotating window. For each orientation of the window the vector field is interpolated into the corners of the window (Figure 1).

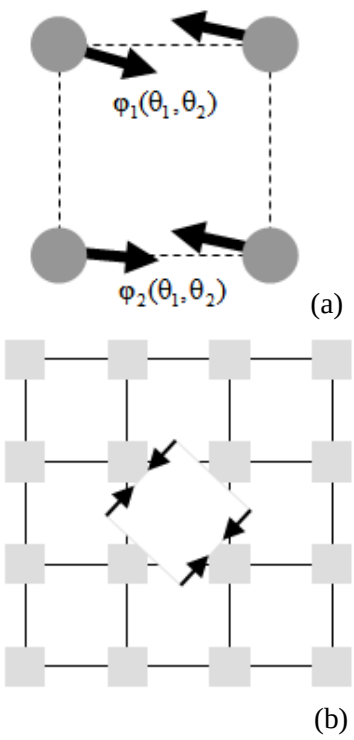


Figure 1. (a) The direction score as the vector alignment measure (b) the rotating window

Next, $\varphi(\theta_1,\theta_2)$ is the deviation of the vectors at the two opposing corners from the direction corresponding to the orientation of the window. The closer the two vectors are to the prescribed direction the greater is $\varphi(\theta_1,\theta_2)$.

The maximum response $\varphi(\theta_1,\theta_2)$ shows how close to the anti-parallel position the vector field is in the locality of the candidate boundary point. In order to construct the membership function $\varphi(\theta_1,\theta_2)$ we use two dimensional interpolation techniques with a few control points. The measure $\varphi(\theta_1,\theta_2):([0,2\pi]\times[0,2\pi]) \to [0,1]$ is constructed on a triangle $\left\{\left(0,\frac{\pi}{2}\right),\left(\frac{\pi}{2},\frac{\pi}{2}\right),\left(\frac{\pi}{2},\pi\right)\right\}$ using interpolating points at the corners and in the middle of each side of the triangle. The following conditions at the interpolating points are used:

$$\varphi(0,\pi) = 1,\ \varphi\left(\frac{\pi}{2},\pi\right) = 0,\ \varphi\left(\frac{\pi}{4},\frac{3\pi}{4}\right) = \beta$$

$$,\ \varphi\left(0,\frac{3\pi}{4}\right) = \alpha,\ \varphi\left(\frac{\pi}{2},\frac{\pi}{2}\right) = 0,\ \varphi\left(\frac{\pi}{2},\frac{3\pi}{4}\right) = 0,$$

where α and β are the design values. Suppose that $\varphi\left(0,\frac{3\pi}{4}\right)=\alpha=0.5$. It means that the measure of position (d), deviating from the ideal position by $\frac{\pi}{4}$, is 50% of the what we assign in the ideal case when the deviation is zero. Furthermore, the six positions define a quadratic polynomial given by

$$\varphi(\theta_1,\theta_2)=a_1\theta_1^2+a_2\theta_1\theta_2+a_3\theta_2^2+a_4\theta_1+a_5\theta_2+a_6$$

Once coefficients a_i are evaluated, $\varphi(\theta_1,\theta_2)$ is extended to $[0,2\pi]\times[0,2\pi]$ symmetrically, thus, becoming a piecewise quadratic function.

Note, that if $\alpha=\beta=\frac{1}{2}$, then $\varphi(\theta_1,\theta_2)$ becomes a bilinear function given by $(1-\frac{2}{\pi}\theta_1)(1-\frac{2}{\pi}\theta_2)$. $\varphi(\theta_1,\theta_2)$ applies to rotating window W_γ (Figure 1) to produce a sequence of two dimensional score vectors $s_\gamma=(\varphi_1,\varphi_2)_\gamma$, where γ is the rotation angle. The vectors are normalized as follows $s_{\gamma,\text{new}}=\frac{s_\gamma}{\max|s_\gamma|}$. Finally, the directional score $s_D=\min_\gamma|s_{\gamma,\text{new}}|$. We assume that the true contour points form large clusters characterized by $s_D>1-\delta$, where δ is an appropriate threshold. Besides the true boundary points must belong to relatively large continuous segments, whereas the false contour points belong to short segments. Consequently, appropriate edge detection methods such as [43-45] can be applied.

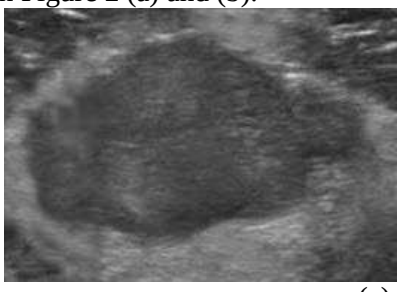
COFFA can be applied in the iterative way as follows. First, the GGVF is applied to the raw gradient field of the edge map ∇f to obtain the vector function $V(x,y)=(u(x,y),v(x,y))$. Next, we measure the COFFA score $s_D(x,y)$ and run an iterative improvement of s_D by obtaining a new the vector function $V(x,y)$.However, using the iterative COFFA-images slows down the algorithm. Besides, this procedure may lead to numerical instabilities. This version of the algorithm calculates s_D only once and uses the corresponding directional score in the subsequent calculations.

7. Dynamic diffusion

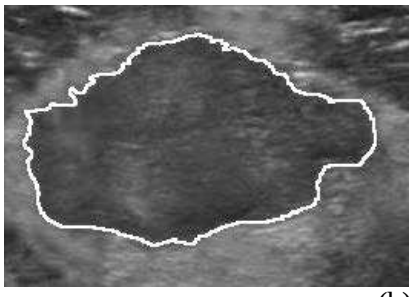
As opposed to the classical GGVF, where the diffusion is defined by a coefficient K in (2)-(3), our model employs a rational function $-\frac{|\nabla f|}{K(|\nabla f|)}$, where K is a monotone Hermite cubic spline, $K>0, K'<0$. The polynomial is generated using significant clusters of the edge map generated by a standard fuzzy C- mean method. Usually, the algorithm automatically generates four clusters corresponding to background, noise, shadows and the boundary of the tumor. The centroid s_i of each cluster is selected as the representative of the cluster and the corresponding Hermite polynomial is constructed so that $K(s_i)=\zeta_i$, where ζ_i are selected manually. The proposed polynomial is the monotone piecewise cubic Fritsch-Carlson spline constructed using the slopes of the secant lines between successive points and adjusting the slopes to ensure monotonicity. After the preprocessing, the evaluation of the interpolated spline is equivalent to that of the standard Hermite spline.

8. Analysis of Sensitivity to Initialization.

This example demonstrates a test performed to analyze the sensitivity to the position of the initial contour. The test tumor and the ground truth are shown in Figure 2 (a) and (b).



(a)



(b)

Figure 2 Testing the sensitivity of the initialization.

The snake has been initialized at the distance d= 5, 12, 32 pixels from the boundary measured by the Hausdorff distance H_1. The results for each initialization are presented in Tables 1, 2 and 3 respectively

Table 1. Accuracy and convergence of GGVF-COFFA snakes, d=5.

Method/ Criterion	Preprocessing		
	GGVF, best K = 0.009,0.01	GGVF with adaptive K	COFFA-GGVF
TP	99.1028	99.1028	99.1905
H_1	2.1225	2.1472	2.1217
H_2	0.3223	0.3199	0.3088
H_3	0.2027	0.2012	0.1942
Snake convergence	Yes	Yes	Yes
GGVF convergence	Yes	Yes	Yes

Table 2. Accuracy and convergence GGVF-COFFA snakes, d=12

Method/ Criterion	Preprocessing		
	GGVF, best K = 0.08	GGVF with adaptive K	COFFA-GGVF
TP	88.8755	99.1026	99.1905
H_1	23.1307	2.1680	2.1410
H_2	1.4630	0.3218	0.3104
H_3	0.9201	0.2024	0.1952
Snake convergence	Yes	Yes	Yes
GGVF convergence	No	Yes	Yes

Table 3. Accuracy and convergence GGVF-COFFA snakes, d=32

Method Criterion	Preprocessing		
	GGVF, best K = 0.09	GGVF with adaptive K	COFFA-GGVF
TP	83.7552	99.1025	99.1903
H_1	29.4658	2.1703	2.1461
H_2	3.2775	0.3236	0.3221
H_3	2.0613	0.2034	0.2026
Snake convergence	Yes	Yes	Yes
GGVF convergence	No	Yes	Yes

When the initial snake is close to the true boundary, the performance of GGVF and COFFA-GGVF is excellent. For d=5, the accuracy is 99% in terms of true positives and 0.3 in terms of H_2. Moreover, GGVF converges. Yet, increasing the distance leads to a drastic decline in the accuracy of GGVF in terms of the true positives (TP) as well as in terms of the max- H_1, average-H_2 and relative-H_3 Hausdorff distance. The noise which was not properly processed by the conventional GGVF attracts the snake leading to huge errors. For d=32 the best performance of GGVF is only 83% in terms of true positives and the maximum deviation from the true contour reaches 29 pixels. However, COFFA-GGVF produces a consistent result of 99% TP. The deviation measured by H_1 is of about 2.14 pixels

Conclusions

The GGVF-snake segmentation is an efficient and well established method of segmentation. The snakes are able to work in different modes which include (but limited to) adaptive diffusion, multiple snakes, region based snakes, shape prior snakes and many other modifications. The recent adaptive COFFA-GGVF snakes show much promise as applied to segmentation of the breast tumors. The method generates the same or even better accuracy with the reference to the standard GGVF method equipped with the best set of parameters found for the particular image.

References

[1] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, *International Journal of Computer Vision*. 4: 321–331, 1988.

[2]. F. Lefebvre, G. Berger, P. Laugier, Automatic detection of the boundary of the calcaneus from ultrasound parametric images using an active contour model; clinical assessment, *IEEE Transactions on Medical Imaging*, 17 (1): 45-52, 1998.

[3]. Y.S. Akgul, C. Kambhamettu, M. Stone, Extraction and tracking of the tongue surface from ultrasound image sequences, International IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, California, 298-303, 1998.

[4]. R. Chung, C.K. Ho, Using 2D active contour models for 3D reconstruction from serial sections. Proceedings of the 13th International IEEE Conference on Pattern Recognition, Vienna, Austria, 849-853, 1996.

[5]. A. Fenster, S. Tong, H.N. Cardinal, C. Blake, D.B. Downey, Three-dimensional ultrasound imaging system for prostate cancer diagnosis and treatment, *IEEE Transactions on Instrumentation and Measurement*, 47 (6): 1439-1447, 1998.

[6]. M.G. Strintzis, I. Kokkinidis, Maximum likelihood motion estimation in ultrasound image sequences, *IEEE Signal Processing Letters*, 4 (6):156-157, 1997

[7]. D.R. Chen, R.F. Chang, W.J. Wu, W.K. Moon, W.L. Wu, 3-D Breast ultrasound segmentation using active contour model, *Ultrasound in Med. & Biology*, 29 (7):1017-1026, 2003

[8]. R.F. Chang, W.J. Wu, W.K. Moon, W. Lee, D.R. Chen, Segmentation of breast tumor in three-dimensional ultrasound images using three-dimensional discrete active contour model, *Ultrasound in Medicine & Biology*, 29 (11): 1571-1581, 2003.

[9] M. Cvancarova, F. Albresgtsen, K. Brabrand, E. Samset, Segmentation of ultrasound images of liver tumors applying snake algorithms and GVF, *International Congress Series*,1281:218-223, 2005.

[10] M. Aleman-Flores, P. Aleman-Flores, L. Alvarez-Leon, M.B. Esteban-Sanchez, R. Fuentes-Pavon, J.M. Santana-Montesdeoca, Computerized ultrasound characterization of breast tumors, *International Congress Series*, 1281: 1063-1068, 2005.

[11] L.D. Cohen, On Active Contour Models and Balloon, CVGIP: *Image Understanding*, 53 (2): 211-218, 1991.

[12] L.D. Cohen, I. Cohen, Finite-element methods for active contour models and balloons for 2-D and 3-D images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15 (11): 1131 – 1147, 1993.

[13] O. Pujol, D. Gil, and P. Radeva, Fundamentals of stop and go active models, *Image Vision and Computing*, 23:681–691,2005.

[14] J. Tang, A multi-direction gvf snake for the segmentation of skin cancer images, *Pattern Recognition*, 42: 1172–1179, 2009.

[15] F. Y. Shih and K. Zhang, Locating object contours in complex background using improved snakes, *Computer Vision and Image Understanding*, 105: 93–98, 2007.

[16] I. Dagher and K. E. Tom, Waterballoons: A hybrid watershed balloon snake segmentation, *Image Vision and Computing*,26: 905–912, 2008.

[17] X. Zhu, P. Zhang, J. Shao, Y. Cheng, Y. Zhang, and J. Bai, A snake-based method for segmentation of intravascular ultrasound images and its in vivo validation. *Ultrasonics*, 51(2): 181–189, 2011.

[18] S.D. Fenster, J.R. Kender, Sectored snakes: evaluating learned energy segmentations, *Transactions in Pattern Analysis and Machine Intelligence*, 23 (9): 1028–1034, 200

[19] M.-A. Charmi, S. Derrode, F. Ghorbel, Fourier-based geometric shape prior for snakes, *Pattern Recognition Letters*, 29: 897–904, 2008.

[20]. C. Chesnaud, Ph. Réfrégier, V. Boulet, Statistical region snake-based segmentation adapted to different physical noise models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 (11): 1145-1157, 1999.

[21] R. Ronfard, Region-based strategies for active contour models, 13 (2): 229–251, 1994.

[22] S.C. Zhu, A. Yuille, Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, *IEEE Transactions in Pattern Analysis and Machine Intelligence*, 18 (9): 884–900, 1996

[23] T. McInerney and D. Terzopoulos, T-snakes: topology adaptive snakes, *Medical Image Analysis*, 4(2): 73-91, 2000.

[24] G. Giraldi, E. Strauss and A. Oliveira, Dual-T-Snakes model for medical imaging segmentation, *Pattern Recognition Letters*, 24: 993–1003, 2003.

[25] H. Delingnette and J. Montagnat, New algorithm for controlling active contours shape

and topology, 6th European Conference on Computer Vision, Ireland, 2: 381-395, 2000.
[26] H. Delingette, J. Montagnat, Shape and topology constraints on parametric active contours, *Computer Vision and Image Understanding*, 83: 140–171, 2001.
[27]. R. Malladi, J. Sethian, B. Vemuri, Shape modeling with front propagation, *IEEE Transactions in Pattern Analysis and Machine Intelligence* , 17 (2): 158–171, 1995.
[28] S. Osher and J. A. Sethian, Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation, *Journal of Computational Physics*, 79: 12-49, 1998.
[29] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *International Journal of Computer Vision*, 22 (1): 61–79, 1997.
[30] K. Siddiqi, Y.B. Lauzie`re, A. Tannenbaum, S.W. Zucker, Area and length minimizing flows for shape segmentation, *IEEE Transactions in Image Processing*, 7 (3): 433–443, 1998.
[31] X. Wang, L. He, W.G. Wee, Deformable contour method: a constrained optimization approach, *International Journal of Computer Vision*, 59 (1): 87–108, 2004.
[32] L. He, Z.Peng, B.Everding, X.Wang, C.Y. Han, K.L. Weiss, W.G. Wee, A comparative study of deformable contour methods on medical image segmentation, *Image and Vision Computing*, 26: 141–163, 2008
[33] C. Xu, D. Pham, J. Prince, Image segmentation using deformable models, in: Handbook of Medical Imaging, 2, Medical Image Processing&Analysis, SPIE Press,129-174, 2007
[34] C. Li, J. Liu, and M. D. Fox, Segmentation of external force field for automatic initialization and splitting of snakes. *Pattern Recognition*, 38(11): 1947–1960, 2005
[35] M. Rochery, I. H. Jermyn, J. Zerubia, Higher order active contours, *International Journal of Computer Vision*, 69 (1): 27-42, 2006.
[36] C. Xu, J.L. Prince, Snakes, shapes, and gradient vector flow, *IEEE Transaction on Image Processing*, 7 (3): 359–369, 1998.
[37] C. Xu, J.L. Prince, Gradient vector flow: A new external force for snakes, Proceedings of the international IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 66 – 71, 1998.
[38] C. Xu, J.L. Prince, Generalized gradient vector flow external forces for active contours, *Signal Processing*, 71 (2): 131-139, 1998.
[39] J. Tang, A multi-direction GVF snake for the segmentation of skin cancer images, Pattern *Recognition*, 42: 1172-1179, 2009
[40] M. Wei, Y. Zhou, M. Wan, A fast snake model based on non-linear diffusion for medical image segmentation, *Computerized Medical Imaging and Graphics*, 28 (3):109–117, 2004.
[41] J. Mille, Narrow band region-based active contours and surfaces for 2d and 3d segmentation, *Computer Vision and Image Understanding*,113: 946–965, 2009
[44] Y. Shang, X. Yang, L. Zhu, R. Deklerck, and E. Nyssen, Region competition based active contour for medical object extraction, *Computerized Medical Imaging and Graphics*, 32(2): 109–117, 2008.
[45] A. K. Jumaat, W. E. Z. W. A. Rahman, A. Ibrahim, and R. Mahmud, Segmentation of masses from breast ultrasound images using parametric active contour algorithm, *Procedia - Social and Behavioral Sciences*, 8: 640 – 647, 2010.
[46] D. Chen, R. Chang, W. Wu, W. Moon, and W. Wu, 3-d breast ultrasound segmentation using active contour model, *Ultrasound in Medicine and Biology*,29(7): 1017–1026, 2003
[47] R.-F. Chang, W.-J. Wu, W. K. Moon, W.-M. Chen, W. Lee, and D.-R. Chen, Segmentation of breast tumor in three-dimensional ultrasound images using three-dimensional discrete active contour model, *Ultrasound in Medicine and Biology*, 29(11): 1571–1581, 2003.
[46] T. McInerney, Sketchsnakes: Sketch-line initialized snakes for efficient interactive medical image segmentation, *Compiterized Medical Imaging and Graphics*, 32(5): 331–352, 2008.
[47] G. Hamarneh and T. Gustavsson, Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images, in Computers in Cardiology, 115 –118, 2010.
[48] C. M. Chen, H. H. S. Lu, and Y. C. Lin, An early vision-based snake model for ultrasound image segmentation, *Ultrasound in Medicine and Biology*, 26(2): 273 – 285, 2000.
[49] M. Mignotte and J. Meunier, A multiscale optimization approach for the dynamic contour-based boundary detection issue, *Computerized Medical Imaging and Graphics*, 25(3): 265 – 275, 2001.
[50] A. Rodtook and S. S. Makhanov, Continuous force field analysis for generalized gradient vector flow field, *Pattern Recognition*, 43(10): 3522 – 3538, 2010.

[51] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, Automated breast cancer detection and classification using ultrasound images: A survey, *Pattern Recognition*, 43(1): 299–317, 2010.
[52] J. S. Lee, Digital image enhancement and noise filtering by using local statistics, *IEEE Transactions in Pattern Analysis and Machine Intelligence*, PAM1-2: 165-168, 1980.
[53]V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, A model for radar images and its application to adaptive digital filtering of multiplicative noise, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4: 157–165, 1982.
[54]. D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, Adaptive restoration of images with speckle, *IEEE on Transactions Acoustic, Speech and Signal Processing*, ASSP-35: pp. 373–383, 1987.
[55] S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, A versatile technique for visual enhancement of medical ultrasound images, *Digital Signal Processing*, 17: 542–560, 2007.
[56] Y. Zhang, R. Sankar, and W. Qian, Boundary delineation in transrectal ultrasound image for prostate cancer, *Computers in Biology and Medicine*, 37: 1591–1599, 2007.
[57] B. Liu, H. D. Cheng, J. Huang, J. Tian, X. Tang, and J. Liu, Probability density difference-based active contour for ultrasound image segmentation, *Pattern Recognition*, 43: 2028–2042, 2007.
[58] P. Perona and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12: 629–639, 1990.
[59] R. N. Czerwinski, D. L. Jones, and W. D. O'Brien, Detection of lines and boundaries in speckle images - application to medical ultrasound, *IEEE Transactions on Medical Imaging*, 18: 126–136, 1999.
[60] A. Lopes, R. Touzi, and E. Nezry, Adaptive speckle filters and scene heterogeneity, *IEEE Transactions on Geoscience and Remote Sensing*, 28: 992–1000, 1990.
[61]. Y.-L. Huang, Y.-R. Jiang, D.-R. Chen, and W. K. Moon, Level set contouring for breast tumor in sonography, *Journal of Digital Imaging*, 20(3): 238–247, 2007.
[62] B. Levienaise-Obadia and A. H. Gee, Adaptive segmentation of ultrasound images, *Image Vision and Computing*, 17(8): 583–588, 1999.
[63] J. Yu and J. Tan, Object density-based image segmentation and its applications in biomedical image analysis. *Computer Methods and Programs in Biomedicine*, 96: 193–204, 2009.
[64] L. Vincent, Morphological grayscale reconstruction in image analysis Applications and efficient algorithms, *IEEE Transactions on Image Processing*,2: 176–201, 1993.
[65] M. Mignotte, J. Meunier. A multiscale optimization approach for the dynamic contour-based boundary detection issue, *Computerized Medical Imaging and Graphics*, 25: 265-275, 2001.
[66]. Nilanjan Ray, Bhabatosh Chanda, Jyotirmay Das, A fast and flexible multiresolution snake with a definite termination criterion, *Pattern Recognition*, 34 : 1483-1490, 2001
[67]. Chucherd, S., A. Rodtook and S.S. Makhanov. Phase portrait analysis for multiresolution generalized gradient vector flow, *IEICE Transactions on Information and Systems*, E93-D (10): 2822-2835, 2010.
[68]. Fritsch F.N. and Carlson R.E., Monotone Piecewise Cubic Interpolation, *SIAM J. of Numerical Analysis*, 17: 238-246, 1980.
[69] Castelmen K.R. , Digital Image Processing, Prentice Hall, 1996.
[70] Chipman H. and Tibshirani R. Hybrid hierarchical clustering with applications to microarray data, *Biostatistics* 2(7): 286–301, 2006.
[71]. G.Naldi, L.Pareschi, G.Toscani, Relaxation schemes for PDEs and applications to fourth order diffusion equations, *Surveys on Mathematics for Industry*, 10: 315-343, 2002.

Optimization of 5-axis machining with application to dental implants

Samart Moodleah, Stanislav S. Makhanov

School of Information and Computer Technology
Sirindhorn International Institute of Technology
Thammasat University, Thailand
samartjob@gmail.com, makhanov@siit.tu.ac.th

Abstract

Machining large complex industrial often requires tens, hundreds of cutter location points and hundreds of hours. That is why reducing the machining time is a one of the most important topics in optimization of CNC codes for five axis milling machines. Maximization of the machining strip employing a flat end is one of the strategies towards this goal. Each position on the surface there exists at least one direction which maximizes the machining strip.

In this paper we propose a new method of constructing a curvilinear tool paths which partly or even entirely align with the vector field of the optimal directions at the points characterized by large curvature. The vector field aligned tool path (VFAP) is based on the curvilinear elliptic grid generation and requires numerical solution of a corresponding system of PDEs. The curves from the numerically generated curvilinear grid are converted into a tool path using the space filling curve (SFC).

The combination of VFAP and SFC has been compared with a standard iso-parametric zigzag tool path and also compared with tool paths generated by MasterCam X5. The numerical experiments demonstrate that the method outperforms the standard iso-parametric zigzag in terms of the length of the tool path and ultimately the machining time.

Five axis machining technology has been widely used in multi area including Dental reproduced, tooth implant. Advances in computer technology now enable cost-effective production of individual pieces. Dental restorations produced with computer assistance have become more common in recent years.

Keywords: Tool path planning, error minimization, space filling curve, dental reproduce CAD/CAM

1. Introduction

Milling machines are programmable mechanisms for cutting complex industrial parts designed in such a way that the cutting tool (cutter) is capable of approaching the desired surface at a given point with a required orientation.

The machine is built of several moving parts designed to establish the required coordinates and orientations of the tool during the cutting process. The movements of these parts are guided by a controller which is fed with the so-called NC program or G-code. The G-code comprises commands carrying three spatial coordinates of the tool-tip and a pair of rotation angles needed to establish the orientation of the tool.

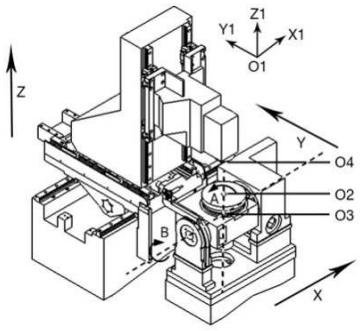


Figure 1. Five axis milling machine with rotation axes of the table (Kasetsart University, Thailand)

In this paper we consider a tool path optimization performed with regard to machining time. In order to reduce the machining time we minimize the total length of the tool path subject to the scallop height constraint. The scallop height constraint is satisfied by evaluating the machining strip and selecting an appropriate offset between the adjacent tool trucks. The maximum machining

strip leads to a larger distance between the tracks. In turn, this reduces the number of tracks and consequently the length of the tool path.

Following the optimal or nearly optimal directions can be combined with rear gouging, global gouging and machine limits constrains. For instance, in [1] the authors generate an accessibility map which is composed of the admissible range of the inclination and yaw angles and the smoothness map based on the derivatives of the tool vectors evaluated at the prescribed cutter location points. The two maps combined with the path propagation algorithm similar to [2]. However, the algorithm also requires an initial path. Besides the smoothness map[1] is not efficient from the kinematics error viewpoint. If the surface has many stationary points (the minimum, the maximum or the saddle points). In this case small variations of the tool vector could lead to large variations of the rotation angles [3]. In [4] this effect is not taken properly into account.

In this paper we propose a curvilinear tool path aligned with the vector field of the optimal directions based on the elliptic grid generation. The vector field aligned path (VFAP) is a compromise between the flexibility of methods based on propagating the initial tool path and the simplicity of the zigzag tool path topology. The techniques are borrowed from computational fluid dynamics where the idea to align the curvilinear grid to the vector field was exploited to improve the accuracy of the numerical solution of PDEs [5]. In our case the grid is aligned with the optimal directions and the tool path represented by a pre-computed vector field.

The grid is generated in the parametric domain and does not require a special initialization. As opposed to that finding the initial tool path [6] is a computationally expensive NP-hard problem although good initializations such as in [1] can lead to a 40% decrease in the tool path length. Nevertheless, the sensitivity of the method to the initial path could be high. Some initial curves may not lead to any improvement in the machining time.

Finally, space-filling curves based on curvilinear grids were introduced in [7]. The proposed method provides considerably faster machining with the reference to the iso-parametric tool path and a tool path generated by MasterCam X5.

2. Inclination of the Tool and Optimal Cutting Direction

Practically, the radius of the tool must not larger than the radius of the curvature in the direction perpendicular to the cutting curve. Changing the tool size to always match with the radius of the cutting curve is may less efficient methodology. As opposed to that, a flat end tool Figure 2 (a) can be inclined dynamically in order to avoid the curvature interference [8],[9] as shown in Figure 2 (b) and Figure 2 (c).

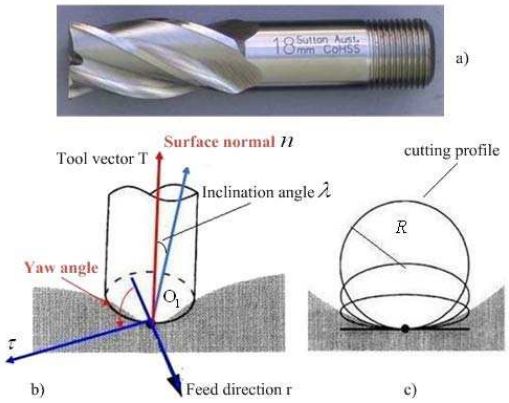


Figure 2. Flat end tool and the tool cutting profile.

The tool orientation is controlled by the inclination angle λ and a yaw angle w. For simplicity in this paper we consider $w = 0$. In this case the minimal inclination angle to avoid the curvature interference is given by [10]

$$\lambda = \sin^{-1}(Rk_{\max}),$$

where R is the tool radius and $k_{\max}$ the maximum curvature at a cutter contact point. The tool vector T is obtained by rotating the normal n towards r (the feed direction) as follows $T = \cos(\lambda)n + \sin(\lambda)r$ (Figure 2 (b)) [11].

To evaluate the machining strip width, the surface cross-section perpendicular to the tool cutting direction r is approximated by a circular arc, for which the radius R is equal to the radius of the curvature of the surface in the r direction. Given the maximum scallop height h, the machined surface error is represented by a virtual circular arc of radius $R-h$ concentric

with the first circular [11]. The machining strip is then obtained by locating the intersections of the effective cutting shape with this virtual circular. The machining strip is used to offset the curvilinear tool paths. The offset is evaluated at the post processing stage when the entire curvilinear grid has been generated.

3. Tool path generation

$S \equiv S(u,v) \equiv (x(u,v), y(u,v), z(u,v))$ denotes the part surface, where u and v are the parametric variables. Consider a set of CC points $\{u_{i,j}, v_{i,j}\}$ arranged as a curvilinear grid. Mathematically, it means that set $\{(u_{i,j}, v_{i,j}),$ $0 \le i \le N_\xi \;\; 0 \le j \le N_\eta\}$ is a discrete analogy of a mapping from the computational region $\Delta = \{0 \le \xi \le N_\xi,\; 0 \le \eta \le N_\eta\}$ onto a parametric region defined in the parametric coordinates *u, v*.

Consider a vector field of optimal directions given in the parametric domain by $r(u,v) \equiv (\alpha(u,v), \beta(u,v)) = r(\xi,\eta)$, $\| r \| \equiv 1$. Now we will align the curvilinear grid $\{u(\xi,\eta), v(\xi,\eta)\}$ and $r(\xi,\eta)$. There are many grid generation methods capable to achieve this goal. In this paper we use one of the simplest classical methods [12]. The required grid satisfies

$$\begin{cases} u_\xi = \;\; \alpha(\xi,\eta) \\ u_\eta = -\beta(\xi,\eta) \end{cases}, \quad \begin{cases} v_\xi = \beta(\xi,\eta) \\ v_\eta = \alpha(\xi,\eta) \end{cases}.$$

Differentiation and summation yields

$$\begin{cases} u_{\xi\xi} + u_{\eta\eta} = \alpha_\xi - \beta_\eta, \\ v_{\xi\xi} + v_{\eta\eta} = \alpha_\eta + \beta_\xi. \end{cases} \tag{1}$$

We augment the system elliptic PDEs (1) with appropriate boundary conditions in accordance with the use of a simple canonical geometry in the parametric space. Namely, if the parametric region in (u,v) is a unit square, the problem is of mapping a rectangle $[0 \le i \le N_\xi] \times [0 \le i \le N_\eta]$ onto the unit square subject to the alignment of the mapping to the given vector field. The required mapping is $\{u(\xi,\eta), v(\xi,\eta)\}$. As far as the problem of tool path generation is concerned, usually the parametric region is a unit square.

The problem can be further simplified further by noting that it is important to follow the optimal directions only where the amplitude of the curvature $k_{\max} = \max|k|$ is high. Only in this case the alignment with the vector field leads to a significant reduction of the tool path. Therefore, the desired vector field is defined by

$$r(u,v) = \begin{cases} (\alpha(u,v), \beta(u,v)), & \text{if } k_{\max} < k', \\ (0,0), & \text{otherwise}, \end{cases} \tag{2}$$

where k' is the prescribed threshold. We will call $r(u,v)$ the thresholded vector field.

Applying the discrete Laplacian yields an iterative second order finite-difference approximation. The corresponding finite-difference equations are solved by the Newton method [12].

Finally, in the case of tool path generation the parametric region can be extended so that the boundary conditions used for the numerical solution are imposed at an artificial boundary. This makes the method very flexible and allows configurations where the tool path crosses the adjacent sides of the parametric rectangle.

4. Numerical Example

In this section the proposed VFAP-SFC method is compared with the iso-parametric zigzags tool path . We also compare with the tool paths generated by MasterCam X5 for flat end tool machining.

4.1. A unimodal surface

The first example demonstrates the efficiency the VFAP with the reference to the traditional iso-parametric tool path method. Consider a unimodal surface (Figure 3 (a)) with an exponential peak along a diagonal line in the parametric domain (*u*,*v*) given by

$$x(u,v) = 100u - 50,$$
$$y(u,v) = 100v - 50,$$
$$z(u,v) = 30e^{-6(v-1.45u+0.25)^2} + 70.$$

Figure 3 (a) shows the test surface. The corresponding vector field is given in Figure 3(b). Since the curvature of the surface is low outside the diagonal peak, we threshold the vector field by using (2) with $k' < 0$ (see Figure 3 (c)). For a flat-end tool with radius 5 mm the final curvilinear grid is shown in Figure 3 (d). The zigzag and SFC tool path are shown in Figure 3 (e) and Figure 3(f) respectively.

This comparison is presented in Table 1. The length of VFAP based on the adaptive grid technology is shorter by 18-40%. For instance, when the maximum allowable scallop height *h*=0.05, the length of the tool path is about 3400mm shorter. The advantage in terms of the machining time is up to 24%. Note that the generating tool path algorithms of the very popular commercial CAD/CAM software MasterCam X5 for *h*=0.25, 0.1 and 0.05 mm are very large compare to proposed method.

Cutting large complex industrial parts with high accuracy may employ tens, hundreds of thousands and even millions of CC points and hundreds of hours. Therefore, the 40% improvement is significant and may amount to long hours of machining. Since the curvilinear grid naturally adapts in the regions of large curvatures, the VFAP requires one large turn where the conventional path takes several turns leading to a larger distance in the angular space.

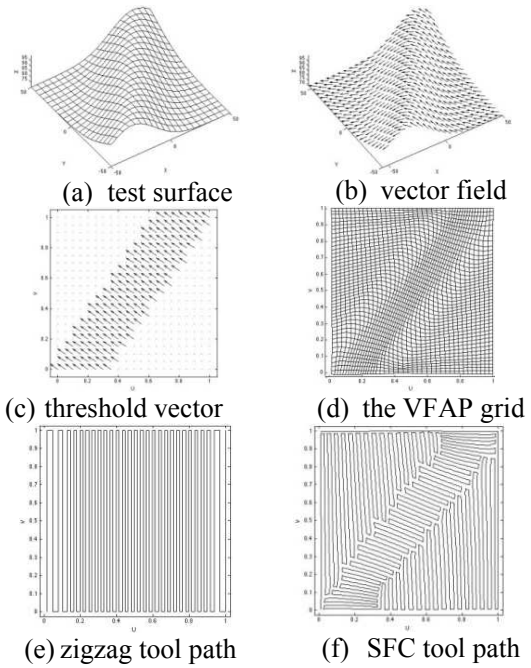

(a) test surface
(b) vector field
(c) threshold vector
(d) the VFAP grid
(e) zigzag tool path
(f) SFC tool path

Figure 3. Numerical result of example 1

4.2. A complicated concave–convex surface with several peaks

Consider a surface shown in Figure 4 (a) given by

$$x(u,v) = 100u - 50,$$
$$y(u,v) = 100v - 50,$$
$$z(u,v) = 11.6e^{-30(v-1.7u+0.3)^2} + 11.6e^{-30(v-1.7u+1.3)^2} + 11.6e^{-30(1.7u-v+0.6)^2} - 33.3v(v-1) + 70.$$

Test surface is shown in Figure 4 (a). The corresponding vector field is displayed in Figure 4 (b). We threshold the vector field with $k' < -0.06$ (see Figure 4 (c)). For a flat-end tool of radius 4 mm, the final curvilinear grid is shown in Figure 4 (d). The zigzag and SFC tool path are displayed in Figure 4 (e) and Figure 4 (f) respectively.

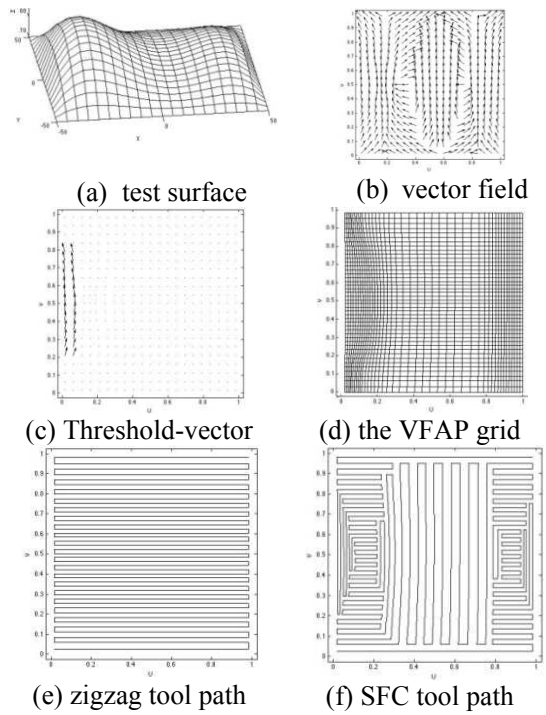

(a) test surface
(b) vector field
(c) Threshold-vector
(d) the VFAP grid
(e) zigzag tool path
(f) SFC tool path

Figure 4. Numerical result of example 2

The Table 2 shows the advantage of the proposed method over the convectional which is 28% in terms of length of the tool path and 38% in terms of the machining. For instance, the maximum allowed scallop of 0.05 mm requires 38 minutes with non optimal tool path and 23 minutes with the optimization.

Table 1. VFAP-SFC vs. conventional tool path surface 1

Tool radius 5 mm		Tool path generation method	Scallop height (mm)		
			0.25	0.10	0.05
Performance	Tool path length (mm)	Master CAM	42508.0	84080.0	131182.0
		Zigzag	3994.9	6052.0	8431.6
		VFAP + SFC	3264.0	3944.3	5079.6
	Machining Time	Zigzag	0:16:08	0:24:55	0:34:55
		VFAP + SFC	0:16:40	0:20:37	0:26:30
Gain	Tool path length (mm)	VFAP + SFC over Zigzag	18.3%	34.8%	39.8%
	Machining Time	VFAP + SFC over Zigzag	-3.3%	17.3%	24.1%

Table 2. VFAP-SFC vs. conventional tool path surface 2

Tool radius 5 mm		Tool path generation method	Scallop height (mm)		
			0.25	0.10	0.05
Performance	Tool path length (mm)	Master CAM	28999.0	57981.0	102749.0
		Zigzag	3317.0	4890.1	6791.8
		VFAP + SFC	2769.3	3656.9	4847.3
	Machining Time	Zigzag	0:18:04	0:26:58	0:38:27
		VFAP + SFC	0:11:31	0:16:42	0:23:45
Gain	Tool path length (mm)	VFAP + SFC over Zigzag	16.5%	25.2%	28.6%
	Machining Time	VFAP + SFC over Zigzag	36.2%	38.1%	38.2%

Five axis machining technology has been widely used in multi area including Dental reproduced, tooth implant. Advances in computer technology now enable cost-effective production of individual pieces. Dental restorations produced with computer assistance have become more common in recent years[13]. It is reported to be 10-50 μm in the marginal area. Thus, the demands of the literature concerning marginal of dental restorations can be reached with this technology; in addition, this production process achieves an industrial standard that does not have to deal with the variations of manually produced prostheses[13].

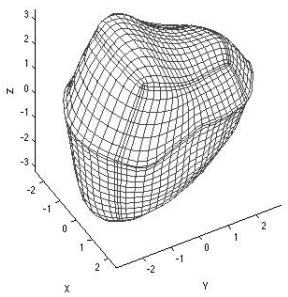


Figure 5. Tooth implant 3D model

As we can see in the examples 4.1 – 4.2, the VFAP technique can be used for generating the tool path correspondent to surface geometry. Consider Figure 4 (a) and Figure 5, the double bell shape is similar to the tooth implant geometry. We can enhance the 5-axis machine ability and combine with the proposed technique to provide the dentist device assist for the tooth implant process.

5. Conclusion

A new method for generation of vector field aligned tool path for five-axis machining has been presented and verified. The method is based on numerical generation of a curvilinear tool path from a corresponding system of PDEs.

The proposed method provides up to 38% decrease of the machining time with regard to the iso-parametric tool path with the dynamic selection of the machining strip while the accuracy is maintain regarding to scallop height constraint.

Acknowledgment

This work was supported by the Thailand Research Fund (TRF), under Royal Golden Jubilee {Ph.D.} Program Grant No. PHD/0330/2551.

Reference

[1] L. L. Li, Y. F. Zhang, H. Y. Li, L. Geng. Generating tool-path with smooth posture change for five-axis sculptured surface machining based on cutter's accessibility map. *International Journal of Advanced Manufacturing Technology*, 53 (5-8):699-709.

[2] Chiou, C.-J. and Lee, Y.-S. A machining potential field approach to tool path generation for multi-axis sculptured surface machining. *Computer-Aided Design*, 34(5):357–371, 2002.

[3] M. Munlin, S. S. Makhanov E.L.J. Bohez. Optimization of rotations of a five-axis milling machine near stationary points. *Computer Aided Design*, 36(12):1117-1128, 2004.

[4] Park, S. C. and Chung, Y. C. Offset tool-path linking for pocket machining. *Computer-Aided Design*, 34(4):299–308, 2002.

[5] J. F. Thompson, B. K. Soni and N. P. Weatherill. *Handbook of Grid Generation.* CRC Press, 1999.

[6] My, C.A., Bohez, E.L.J and Makhanov, S.S. Critical point analysis of 3D vector field for 5-axis tool path optimization. *In Proceedings of the 4th Asian Conference on Industrial Automation and Robotics, ACIAR,* 2005.

[7] Anotaipaiboon W., Makhanov S.S. Curvilinear space-filling curves for five-axis machining. *CAD Computer Aided Design*, 40(3):350-367, 2008.

[8] A. Rao, R. Sarma. On local gouging in five-axis sculptured surface machining using flat end tools. *Computer Aided Design*, 32 :409-420, 2000.

[9] J.H. Yoon. Tool tip gouging avoidance and optimal tool positioning for 5-axis sculptured surface machining. *International Journal of Production Research*, 41:2125-2142, 2003.

[10] W. Anotaipaiboon, S.S. Makhanov. Tool path generation for five-axis NC machining using space-filling curves. *In Proceedings of The Third Asian Conference on Industrial Automation and Robotics*,1:23-28, 2003.

[11] S.S. Makhanov, W. Anotaipaiboon. *Advanced Numerical Methods to Optimize Cutting Operations of Five-Axis Milling Machining*, Springer, 2007.

[12] Castillo J.E and Otto J.S. Numerical Techniques for the Transformation to an Orthogonal Coordinate System Alighned with a Vector Field. *Computers and Mathematics with Applications*, 40: 523-535, 2000.

[13] F. Beuer, J. Schweiger and D. Edelhoff. Digital dentistry: an overview of recent developments for CAD/CAM generated restorations. *British Dental Journal*, 204(9), 2008.

A Study of Retinal Blood Vessel Clustering for Finding the Main Vessel Convergence

[1]**Pongsate Tangseng** [2]**Nittaya Muangnak**
[3]**Pakinee Aimmanee** [4]**Stanislav S. Makhanov** [5]**Bunyarit Uyyanonvara**
Sirindhorn International Institute of Technology
Thammasat University, Thailand
Email: [1]**pongsate1@gmail.com** [2]**csnnym@ku.ac.th**
[3]**pakinee@siit.tu.ac.th** [4]**stanislav@siit.tu.ac.th** [5]**bunyarit@siit.tu.ac.th**

Abstract

This paper proposes a new algorithm for clustering the main blood vessels in retinal images. The new algorithm involves detecting main vessels and clustering. With an assumption that the location of optic disc (OD) is the centroid of the main branches of vessels, we evaluate our algorithm by using accuracy of the centroid of vessels clusters obtained from the experiment being inside of the area of the optic disc. Our new algorithm can obtain the accuracy as high as 92.59%.

Keywords: Data clustering, retinal blood vessel clustering, main vessel convergence

1 Introduction

Data clustering is a process of grouping similar items based on a similarity of patterns and is often one of the first step in data mining analysis [1]. Typical clustering process involves the following steps: pattern representation, similarity measurement, clustering or grouping, data abstraction, and assessment of output.

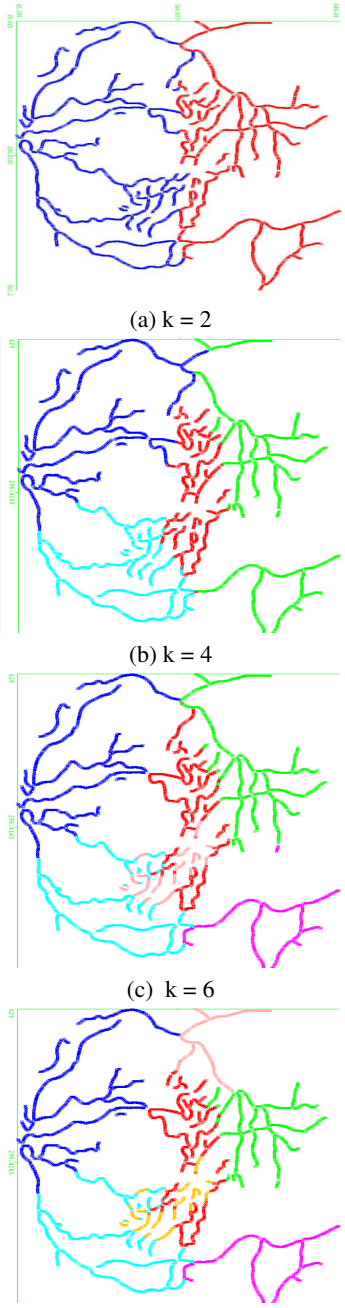
Many clustering algorithms such as Fuzzy c-means [2], Gaussian(EM) [3], Quality threshold [4] and K-mean have been proposed. However, each algorithm has its own merits and demerits and cannot work for all real situations. For example, a well known clustering technique, K-mean clustering [5], is used to classify the series of data items into k clusters for which the within cluster sum of squares is minimum. An example of clusters resulting from K-mean where k equals to 2, 4, 6 and 7 are shown in Figure 1.

However, a major drawbacks of the k means is that the number of clusters must be predefined. Since in this work, we would like to reconstruct main branches of vessels from the list of broken vessels which are obtained from vessel segmentation of the original retina image and the number of clusters (main branches of vessels) are normally unknown, K-mean is not a good choice for our work.

Grouping vessels that are relatively close and have similar properties such as intensity and thickness are key concepts in our clustering algorithm. The purposed method aims to increase the correctness of vessel clustering, which is used to find the location of OD. In the next section, the proposed method is explained in more detail. Section 3 describes the collection of data that is used in the experiment, setting of parameters, and evaluation scheme. Section 4 shows results and discussion and finally the last section summarizes this work.

2 Proposed Method

Our proposed vessel clustering algorithm is described in Algorithm 1. Figure 3 - Figure 7 show the results obtaining from each step in the algorithm. In Figure 2a, retinal blood vessels are extracted using the efficient vessel detection approach based upon using wavelets and edge location refinement [6]. The useful vessel data such as vessels center line, diameter, and length are also obtained from the process of vessel extraction. Figure 2b shows the results after the vessel extraction and segmentation. Since we only care vessels that are relatively close, we also ignore vessels that are too far away from the rest. Figure 3 shows an example of results after removing outbound vessels. In this figure, there are several tiny lines which come from noise in the image. These tiny vessels are then removed from the vessels. Figure 4 shows the example of results after tiny vessels are removed. Then we group vessels that are close. In this step, there are still many clusters. Figure 5 and Figure 6 show the results after we cluster the vessels and



(a) k = 2

(b) k = 4

(c) k = 6

(d) k = 7

Figure 1: Examples of k=2,4,6 and 7 clusters obtained from k-mean clustering of blood vessels at different k's

select only the first top five vessels that are the thickest. Figure 7 we re-merge vessels to get the final cluster.

3 Experiment and Evaluation Results

We use the collection of images collected from the STructured Analysis of the Retina (STARE) research project [7] in our experiment. The collection comprises 81 color images of size

Algorithm 1 Algorithm for finding the clusters of vessels

Require: 1. Vessel data including vessel number, x,y position, thickness and intensity of green channel (G) of each point from vessel extraction. 2. Original image
Let $S \leftarrow$ vessel data; where $Si \leftarrow$ (vessel number, x, y, thickness, G)
$img \leftarrow$ green channel of original image

Step 1: Define high contrast vessels using Difference in Green (DG)
for all $Si, \in S$ **do**
  $window \leftarrow$ img{x-thickness, x+thickness}{y-thickness, y+thickness}
  $thresh \leftarrow$ GRAYTHRESH(window)
  $meanBG \leftarrow$ MEAN(window¿thresh)
  $DG \leftarrow$ G of Si - meanBG
end for

Step 2: Remove low contrast vessels
for all $numberOfVessel$ **do**
  $vesselDG \leftarrow$ mean of DG in a vessel
end for
Find mean and standard deviation of vesselDG
Remove vessels that vesselDG $< param1$

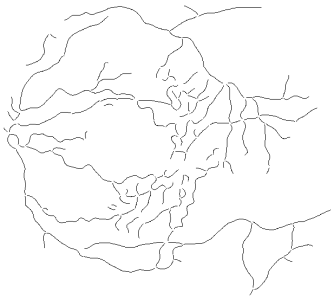
Step 3: Remove short vessels
for all numberOfVessel **do**
  $V \leftarrow$ S; where vessel number = i
  Calculate distance between every 2 points in the V
  Create a weighted complete graph G where each v in V are nodes and weight of edge between each 2 v are distance between them.
  Find minimum spanning tree MST of G
  vessels length is summation of weight in MST
end for
Find mean and standard of vessels length
Remove vessels that vessels length $< param2$

Step 4: Group adjacency vessels
Find distance between vessels
Group vessels that the distance between them is less than $param3$. The group of vessels called cluster.

Step 5: Find length and thickness of clusters
for all $clusters$ **do**
  Create a weighted complete graph G by consider vessels in cluster as a node and distance between vessel is a weight of edge between nodes
  Find minimum spanning tree MST of G
  clusters length is summation of weight in MST
  clusters thickness is summation of thickness of every point in the cluster
end for

Step 6: Select 5 thickest clusters & group adjacency clusters
Find distance between clusters
Group clusters that the distance between them is less than $param4$. The group of clusters is final answer.
return $finalClusters$

(a) An original image



(b) Segmented result

Figure 2: Segmented retinal blood vessel result prior to step 1

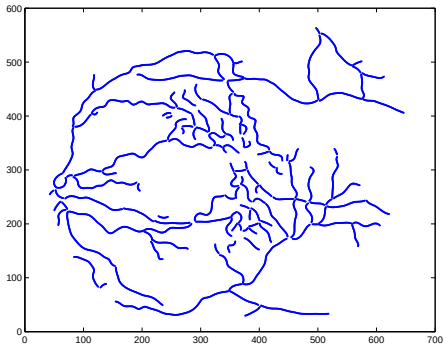


Figure 3: After delete outbound Vessels of DG in step 1

700x605. These images can be classified into two groups: healthy and unhealthy. The numbers of images in the healthy and unhealthy groups are 50 and 31, respectively. For the healthy group, the images are taken from the normal retina. The optic disc boundary and vessels in this set are clearly detected. On the other hand, the optic disc boundary and vessels in the unhealthy group are faint. The boundary of OD and blood vessels are unclear and having close intensity with the background.

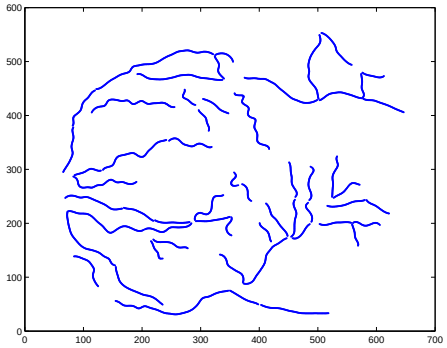


Figure 4: After delete short Vessels in step 3

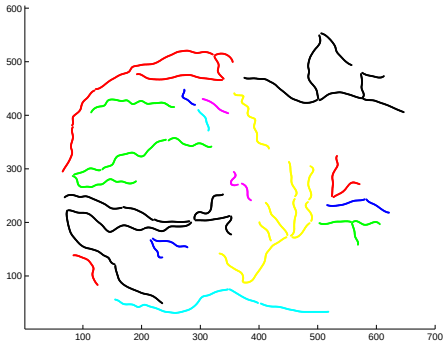


Figure 5: All possible Clusters from step 4

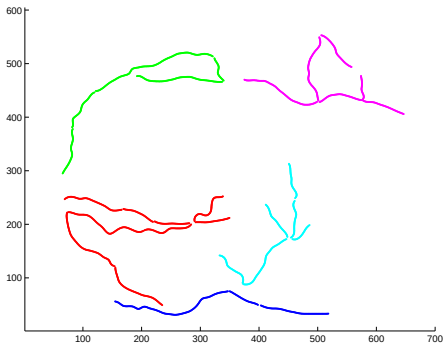


Figure 6: Top five thick Clusters in step 6

First, an image resource was collected from the STructured Analysis of the Retina (STARE) research project [7]. The eighty-one color images are 700x605 pixels in size, and were divided into two groups which are 31 images of healthy retinas and 50 images of diseased retinas. In the images of a healthy group, a clearly identified

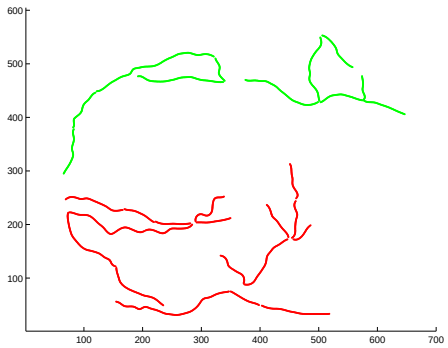
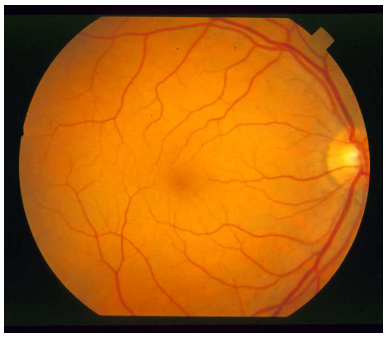
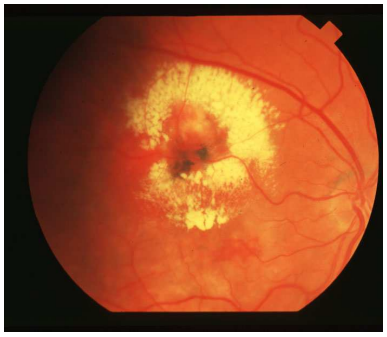


Figure 7: Final Clusters in after step 6

optic nerve has been defined by property contribution of the nerve such as shape, intensity, size, and convergence; meanwhile, the diseased retinal images have been shown anomaly in retina particles that caused the absent of the nerve. The latter makes the retinal vessel detection difficult. Examples of images which are healthy and unhealthy retina are shown in Figure 8a and Figure 8b.



(a) A healthy retinal image

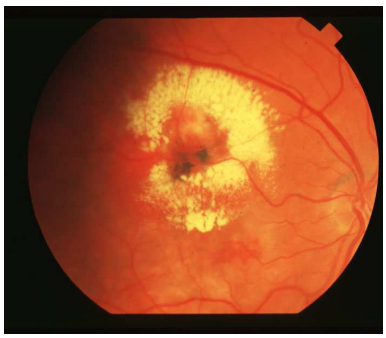


(b) A unhealthy retinal image

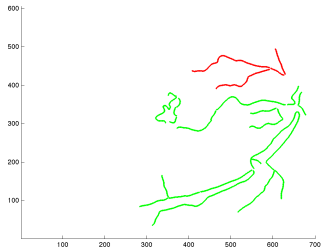
Figure 8: Healthy and unhealthy retinal images

In our experiment, we set the threshold of the difference in green (DG) value be the difference between mean and standard deviation value of DG value, the distance between two vessels to be considered to be in the same cluster be 10, and the length of the vessels to be ignored be the difference between mean and normalized standard deviation value of vessel length where the normalized standard deviation is defined by the standard deviation divided by a constant value.

The selected results obtained from using the proposed methods are shown below in Figure 9 - Figure 12.



(a) original image-1



(b) vessel cluster result-1

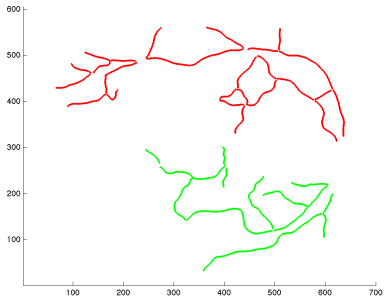
Figure 9: Experimented Results#1

Using the assumption of that the point that is closest to all cluster is the center of the OD [8], we evaluate our vessel clustering [9] by finding the accuracy of getting the location of OD. The accuracy is calculated by finding the ratio of the total number of images whose centroid obtained from the clustering algorithm is inside of the OD to the total number of images. We find that the accuracy of healthy group and unhealthy group are 96.77% and 90.00% , respectively.

4 Conclusion

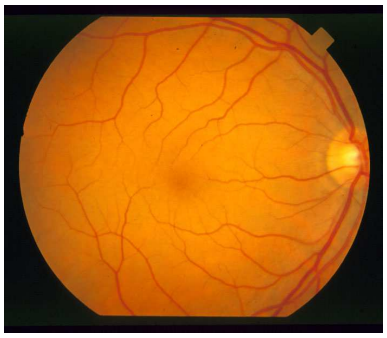
In this paper we proposed a new clustering algorithm which is designed specifically to group the main blood vessels into clusters. These clusters represent main branches of blood vessels. The proposed method uses length, intensity, and thickness of the vessels to detect main vessels

(a) original image-2

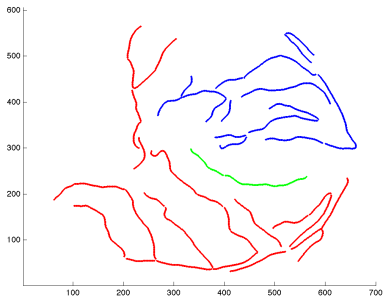


(b) vessel cluster result-2

Figure 10: Experimented Results#2



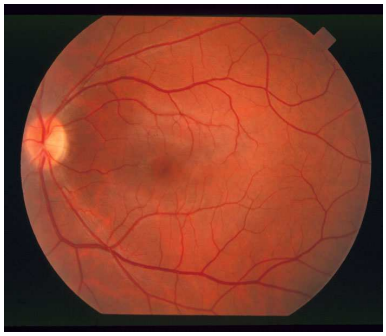
(a) original image-3



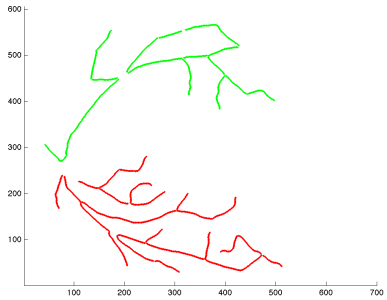
(b) vessel cluster result-3

Figure 11: Experimented Results#3

and uses distances between vessels to cluster two vessels or clusters that are close together. With



(a) original image-4



(b) vessel cluster result-4

Figure 12: Experimented Results#4

assumption that the centroid of the cluster is the location of OD, we find that the accuracy of the algorithm in getting the location of the OD is 92.59%.

Acknowledgment

We acknowledge STructured Analysis of the Retina (STARE) research project for useful standard image database. Thanks to National Research University Project of Thailand Office of Higher Education Commission and Thai Royal Government Scholarship for funding this project.

References

[1] Jain A.K. Murty M.N. and Flynn P.J. Data clustering: A review. *ACM Computing Surveys*, 31, No.3:264–323, September 1999.

[2] Yannis A. Tolias and Stavros M. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, 17, No.2:263–273, April 1998.

[3] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[4] N. Choudhury, R. Sarmah, and S. Sarma. A

modified qt-clustering algorithm over gene expression data. In *Recent Advances in Information Technology (RAIT), 2012 1st International Conference on*, pages 542–547, 2012.

[5] Madhuri A. Tayal and Raghuwanshi M.M. Review on various clustering methods for the image data. *Journal of Emerging Trends in Computing and Information Sciences*, 2, Special Issue:34–38, 2010.

[6] Bankhead P. Scholfield C.N McGeown J.G. and Curtis T.M. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLoS ONE*, 7, Issue 3:1–11, 2012.

[7] STARE project website. http://www.ces.clemson.edu/ ahoover/stare, July 2003. [online].

[8] Bornschlegel T., Muangnak N., Aimmanee P., Makhanov S. S., and Uyyanonvara B. Finding curvess convergence using delaunay search. In *In Proceedings of The First ASIAN Conference on Information Systems (ACIS 2012) [CD-ROM]*, December, 2012.

[9] Maria Halkidi, Yannis Batistakis, and Michalis Vazirgiannis. On clustering validation techniques. *Journal of Intelligent Information Systems*, 17:107–145, 2001.

Comparative Study of Initial Seed Point Required Approaches with the use of Vessel Distance Hint for OD Boundary Detection

[1]Nabwan Prukpaiboon [2]Sarunkorn Chotvijit
[3]Nittaya Muangnak [4]Pakinee Aimmanee
Sirindhorn International Institute of Technology
Thammasat University, Thailand
Email: [1]nabwan.p@gmail.com [2]sarunkorn.cho@gmail.com
[3]csnnym@ku.ac.th [4]pakinee@siit.tu.ac.th

Abstract

This paper presents a comparative study for semi-automated detection of optic disk boundary in fundus images. The study implements and measures performances of region growing, thresholding, and gradient vector flow snake (GVF-snake) approaches. The data set tested in this experiment consists of 91 fundus images, 60 clear images, and 31 faint images. The comparative results show that the thresholding method can estimate the disk boundaries of 91 test images closer to the groundtruth with the better success rate when compared to those obtained by gradient vector flow snake (GVF-snake) and region growing.

Keywords: Boundary detection, region growing, thresholding, GVF-Snake

1 Introduction

Finding optic disc boundary is very important for the medical purpose in the way to detect the abnormality of the patients who suffer from the disease which related to eyes or retinas. The advantages that a system for automatically detect early signs of this disease would provide have been widely studied and experienced positively by experts[1]. In this sense, the optic disc plays an important role in developing automated diagnosis expert systems for diabetic retinopathy as its segmentation is a key preprocessing component in many algorithms designed to identify other fundus features[2]. The relatively distance between the optic disc and the fovea, can be used to help estimate the location of the latter[3]. On the other hand, to segment the vascular tree, blood vessel tracking methods need an initial seed vessel point. For this, pixels of vessels within the optic disc or in its vicinity have been used[4],[5].

Optic disc segmentation is also relevant for automated diagnosis of other ophthalmic pathologies. The most well-known is Glaucoma. It is the second most common cause of blindness worldwide[6]. Image processing is a major tool that has always been chosen to solve a problem in automated ocular treatment and has been introduced to ophthalmology along with robotic surgery [7]. Biomedical engineer in cooperation with computer scientist have developed numbers of approach and algorithms using to increase efficiency rate and reduce chance of error occurrence. With the use of image processing, cataract on optic disc can be easily discovered and targeted by suitable approach and algorithm for the diagnosis of diabetic retinopathy detection[8].

In the literature very few works have been reported on locating the optic disc boundary in retinal images without the initial seed point of the optic disc. Most of the methods only locate the optic disc by using approach such as shape detection or high brightness detection in which they produce different rates of accuracy with some particular drawbacks.[3],[4],[9],[8] These techniques, however, usually fail on pathological images, where not only the optic disc that is characterized as round shape or evaluated brightness, but also other regions of fundus that has similar patterns in which it becomes harder to track the right optic disc and later originate the center point. The technique that proposed color space for optic disc detection and locate the optic disc boundary accurately using an automatically initialized snake. It then apply the used of snake and hand-labelled images to compare the performance of new color morphology methods, but the key problem showed is a comparison of color spaces for performing color morphology to remove blood vessels use for boundary location[10]. S. Sekhar et al.[11] proposed the

method to locate of optic disk by means of using the morphological operations and by using the Hough transform by applying the two main processes, dilation and erosion, to find the optic disc. However, the authors mainly focused on locating the optic disc and did not address the original point of the optic disc which is hard to extend for further developments. Likewise, in[12] a detection in the curves and lines is required a limited data set and may prohibit when dealing with huge amount of datas and sometimes in multidimensional array alike retinal images.

In this work, the boundary of the optic disc is detected by exploring three initial-seed-point required techniques such as region growing, thresholding, and GVF active contour. The initial point of these approach is chosen specially by using shortest distance to all main vessels in the image, The three approaches are applied and implemented for the comparative result in this paper. In addition, we also compare our approches with the others which require no initial seed points to develop a comparative study. Benefits of this project will be a fundamental resources for those who wished to developed an application in ophthalmology field and to detect the possibilities of glaucoma in diabetic patient.

2 Background

Digital image processing, by definition, is the use of complex computer algorithms to create, process, or modify the input digital images, and get modified digital images or some characteristics or parameters related to the digital image in return. The digital image is generally a numerical representation of a two-dimensional image, and is mostly referred to as a raster or a bitmap image. A raster image is a two-dimensional array of pixels, the smallest addressable element that represents the image on the screen. The pixel is a physical point, and it contains 8-bit value of intensity of three different color: red, green, and blue. The three colors are actually three dots which will incorporate into a point-pixel.

There are various techniques in image processing that can be applied in order to detect, modify, or segment the images i.e., extracting the expected features. The basics of techniques applied in this study are described below.

2.1 Region Growing

Region Growing is a simple region-based image segmentation method used to partition an image into regions which, in this case, is applied to extract the optic disc from fundus images. The basic explanation of the mechanism of Region Growing is that it examines neighboring pixels of initial *seed points* and determines whether the pixel neighbors should be added to the region using the pre-defined criteria. The process is then iterated until the criteria is no longer satisfied, in the same manner as in general data clustering algorithms. This is also classified as a pixel-based image segmentation method since it involves the selection of initial seed points.

2.2 Thresholding

Thresholding is a technique used in image segmentation, one of the processes in digital image processing, to convert the image into binary image in which the value of each pixel is either 0 or 1, represented by the colors of black and white. The conversion process is to compare each pixel value of the original image to that of a thresholding value, which can be automatically found, often the mean value, or from some threshold selection algorithm. If the pixel value is greater than the thresholding value, assuming it is an object which should be brighter than its background, then it is set to 1, otherwise it may be the background, and is set to 0.

2.2.1 GVF Snake

Gradient vector flow snake or GVF Snake is an extension of the well known method snakes or active contours. The difference between traditional snakes and GVF Snake is that, GVF Snake converge to boundary concavities and they do not need to be initialized close to the boundary. The core step of Snake algorithm is to minimize the sum of internal and external energy, which both are represented in the form of vectors. Therefore, the external energy and internal energy are both supposed to minimal when the snake is at the object boundary position. In this study, we have tried using GVF Snake to find optic disc boundary.

3 Methodology

3.1 Approach

Fundus or retinal images are often taken by using fundus cameras in hospitals and research centers. We have obtained many free retinal databases from the internet which will come into use during the course of the experiment. The resolution of these images depends upon the quality of the camera used. Most of the new cameras used today produce images of quality that is enough for image processing. This experiment uses the data set of 91 Retinopathy of Prematurity(ROP) images, which is classified into 60 clear images and 31 faint images.[13]

In this study, the experiment flow is designed as indicated in the figure 1 on page 2. The detail explanation of each process is given below:

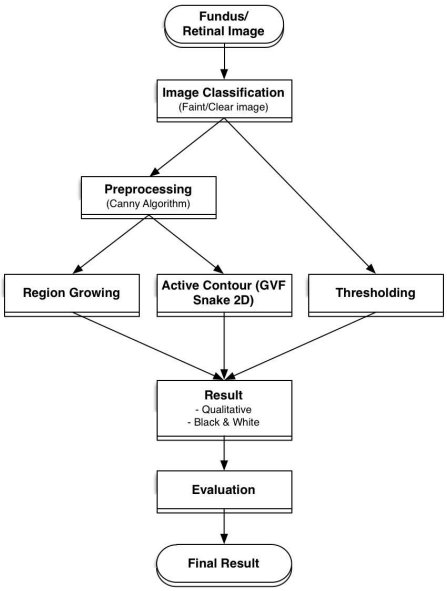


Figure 1: Methodology of study

3.1.1 Image Classification

In this study, since quality and clarity of images has significant effect on performance and accuracy of many algorithms, all images that are tested are manually classified into 2 classes, clear images, and faint images. This classification should also point out each approaches drawbacks regards to quality of source images.

3.1.2 Preprocessing

Fundus images taken by the fundus cameras usually contain noises, illumination spot, or non-uniform features with the brightness centering on the optic disc and decreasing outwards to other regions. This often makes it complicated to extract optic disc from the images as they usually have similar color to the background. Therefore, preprocessing is needed for the ease of feature extraction. The common characteristics of the preprocessing technique used in this study is to apply *Canny operator* to intensify the contrast between the areas of interest and the background. The techniques are often performed on the green channel of RGB color images, which exhibits the area of interest better in terms of contrast than the other channels.

3.1.3 Applying Techniques

In this phase of the study, each fundus image in this experiment are applied with all three approaches. For all approaches, initial seed point is chosen to be the point in the image where the sum of the distance from it to all main vessels are minimum.

- *Region Growing:* Firstly, Canny operator is applied to remove noise in the image, and to enhance the edge of optic disc, making it more suitable for region growing function. Then, apply region growing on the result image.

- *Thresholding:* For each point in the image, we compute shortest distance from it to all main vessels. To get the boundary, we include all points that such a distance is less than the threshold value to be the boundary of the OD. To select the threshold for an image, we find the average of the short distance from each pixel to the main vessels and also find the standard deviation. The threshold is then set to be difference between average value and the standard deviation.

- *Active Contour(GVF Snake):* Firstly, we apply Canny operator to remove noise and to enhance the edge of optic disc, Then, apply the GVF Snake function.

3.1.4 Evaluation and Results

In this study, evaluation was conducted using pixel based technique, in which each pixel of the detection result is compared to the groundtruth image. From the experiment, the number of

pixels that are true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) can be obtained.

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$PositivePredictiveValue = \frac{TP}{TP + FP} \quad (2)$$

This study uses Sensitivity and Positive Predictive Value(PPV) as the performance measure. *Sensitivity* 1 is the measure of the proportion of positives which were correctly detected, while *Positive Predictive Value(PPV)* 2 is the measure of the proportion of correct data from the whole retrieved data.

4 Result & Discussion

From the data set of 60 clear images and 31 faint images, the experiment result is as below:

Region Growing approach gives the result of 44.67% sensitivity and 26.69% PPV for clear images, while giving 74.10% sensitivity and 3.66% PPV for faint images. Thresholding approach gives the result of 62.43% sensitivity and 68.19% PPV for clear images which both are the highest among all approaches, while giving 47.71% sensitivity and 49.12% PPV for faint images. GVF Snake 2D approaches gives 55.93% sensitivity and 18.51% PPV for clear images, while giving 40.68% sensitivity and 9.33% PPV for faint images (see table 1).

Sensitivity of clear ROP image usually provide higher percentage than the faint image. However, for Region Growing, the sensitivity rate of faint image illustrates better result than other approaches. This is because after applying the Canny function, the clearness of the edge sometimes may not explicit enough for the approach to grow only within the area of optic disc boundary. It therefore overgrow the whole retinal image along with the pixel that has similar color intensity value as the region inside the optic disc which finally include all the area of optic disc and result in high percentage of sensitivity.

With a good hint from the shortest distance to the vessels, we can see that most initial seed point required methods work noticeably better than without seed-point required methods (Morphological Operator and Circular Hough Transform) for both Sensitivity and PPV. This implies that the point what gives the minimum sum of shortest distance from it to the main vessels that we used as the seed of the OD is quite useful in getting the location of the OD.

	60 Clear Images		**31 Faint Images**	
	Sensitivity	PPV	Sensitivity	PPV
Implemented in this study				
Region Growing	46.67%	26.69%	74.10%	3.66%
Thresholdding	62.43%	68.19%	47.71%	49.12%
GVF Snake 2D(Active Contour)	55.93%	18.51%	40.68%	9.33%
Result from other study [13]				
Morphological Operator [13]	17.91%	13.12%	10.54%	9.37%
Circular Hough Transform [13]	20.23%	20.11%	9.43%	8.20%

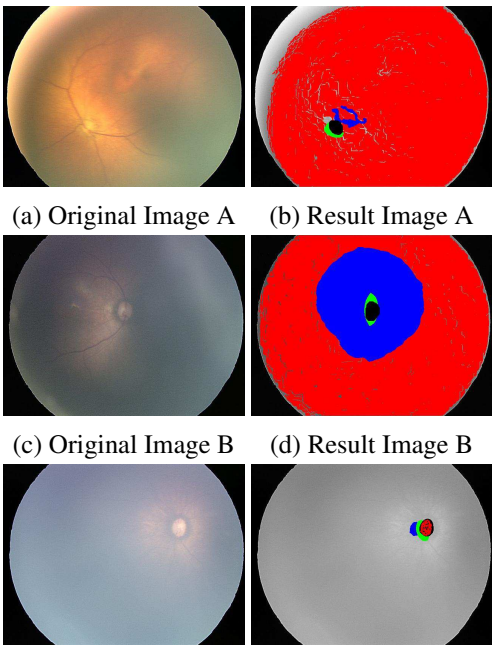
Table 1: Average precision and recall

According to the result, it is obviously seen that thresholding approach gives relatively high PPV not only in clear images but also in low quality images. On the other hand, region growing gives quite high sensitivity while has relatively poor PPV compares to other approaches.

This can be explained that when the edge of the OD is unclear, the region growing approch grows from the inside of the OD and also spurts out to the background resulting very poor PPV. The situation is better when the boundary of the OD is clear. The same scenerio happens with GVF snake, without clear boundary the GVF snake also grows freely to outside.

(See image *Result Image B* in figure Figure 2

on page 4)



(a) Original Image A (b) Result Image A

(c) Original Image B (d) Result Image B

(e) Original Image C (f) Result Image C

Figure 2: Results on sample faint images: Red: Region Growing; Green: Thresholding; Blue: GVF Snake 2D; Black: Optic Disc(OD) boundary

5 Conclusion

In this work, we experiment on three approaches(region growing, vessel distance thresholding, and the GVF snake) of OD detection which requires the initial seed input. We compare the results with other approach which requires no seed point. In our experiment, we pick the seed point of the three methods to be the point whose shorest distance sum to the main vessels in the image is minimum. We found that thresholding approach seems to performs best in both clear and faint images compares to other approaches regards to highest sensitivity and PPV in clear images and the highest PPV in faint images. The major drawbacks that region growing and active contour approach have is that under the constraint of low quality images, the algorithm will overgrow the region (which often enclose the whole optic disk) , resulting in relatively low PPV.

It is observed that most result of thresholding approaches gives the result that is very closed to the actual optic disc, therefore, this approach could be a base resource for development in various detection software that requires optic disc detection.

References

[1] N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan, and I. J. Constable. Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research*, 25:99–127, 2006.

[2] Arturo Aquino, Manuel Emilio Gegúndez-Arias, and Diego Marín. Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *IEEE Transaction on Medical Imaging*, pages 1860–1869, November 2010.

[3] C. Sinthanayothin, J.F. Boyce, H.L. Cook, and T.H. Williamson. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *British Journal of Ophthalmology*, 83:902–910, 1999.

[4] H. Li and O. Chutatape. Automatic location of optic disc in retinal images. In *Proceedings of IEEE International Conference on Image Processing*, pages 837–840, 2001.

[5] L. Gagnon, M. Lalonde, and M. Beaulieu. Procedure to detect anatomical structures in optical fundus images. In *Proceedings of Conference Medical Imaging*, pages 218–1225, San Diego, 2001.

[6] H. A. Quigley and A. T. Broman. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology*, 90:262–267, 2006.

[7] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice-Hall, 2002.

[8] T. Walter, J. C. Klein, P. Massin, and A. Erginay. A contribution of image processing to the diagnosis of diabetic retinopathy—detection of exudates in color fundus images of the human retina. *IEEE Transaction on Medical Imaging*, 21(10):1236–1243, 2002.

[9] M. Lalonde, M. Beaulieu, and L. Gagnon. Fast and robust optic disk detection using

pyramidal decomposition and hausdorff-based template matching. *IEEE Transaction on Medical Imaging*, 20(11):1193–1200, 2001.

[10] Alireza Osareh, Majid Mirmehdi, Barry Thomas, and Richard Markham. Colour morphology and snakes for optic disc localisation.

[11] S. Sekhar, W. Al-Nuaimy, and A.K. Nandi. Automated localisation of retinal optic disc using hough transform. In *The 5th IEEE international symposium on biomedical imaging*, pages 1577–1580, 2008.

[12] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in picture. *Communications of the ACM*, 15:11–15, 1972.

[13] C. Duanggate, B. Uyyanonvara, S.S. Makhanov, S. Barman, and T. Williamson. Parameter-free optic disc detection. *Computerized Medical Imaging and Graphics*, 35(1):51–63, 2011.

Top-view Based People Counting Using Mixture of Depth and Color Information

Chonthisa Wateosot
Faculty of Engineering, Princess of Naradhiwas University
Narathiwat, Thailand, 96000
E-mail: chonthisa.w@pnu.ac.th

Nikom Suvonvorn
Faculty of Engineering, Prince of Songkla University
Hatyai, Songkhla, Thailand, 90112
E-mail: kom@coe.psu.ac.th

ABSTRACT

People counting is very important for some applications, such as, surveillance, event organization, flow control application and etc. In this paper, we propose a novel technique for people counting using the mixture of color and depth images from top-view camera. The U-disparity as depth image projection is introduced in order to increase the accuracy of counting number. Our method can count accurately at 93% in real-time.

KEY WORDS

People counting, Color U-disparity, RGB-D image

1. Introduction

People counting is an important technique that can be used in many applications, such as, tourists flow estimation, retail management, event organization, security application, marketing research pedestrian traffic management and etc. Actually, many researches are done in this field [1], mostly using the motion and color features through object tracking and counting. However, one of the most difficult problems is the people occlusions when considering it at the large entrance gates, such as, supper store, hall or building entrance. For example, a group of people working in the counting area at the same time being very close to each other can produce the partial occlusion. In this situation, only the motion and color properties cannot deal with the problem. However, we believe that the depth information with the corresponding color pattern give us the precious information to solve this problem with a minimum cost and better accuracy.

In this paper, we propose a technique for people counting using the mixture of depth and color images from top-view camera. The key idea is that people will be detected and segmented from depth image instead of color image. And then, using both color and depth images in the tracking process. Based on this idea, the Color U-disparity, defined as an intermediate transformed image space, is applied for people segmentation and tracking. Then, the counting process is applied.

We present our paper in four sections as the following: related work, our proposed system in detail, experimentation result, and conclusion respectively.

2. Related work

Many works has been done by researchers for the people counting method. Albiol et al [1] use the direction of people in the counting process. The gradient stack is applied for estimating the background. The direction of the optical flow is obtained from the horizontal projection of the gradient image. This method cannot count several people when there is an occlusion between them. Xi Zhao et al [2] count the number of people based on face detection, tracking and trajectory classification. The face-tracking method combines a new scale invariant Kalman filter with kernel based tracking algorithm. The accuracy rate is up to 93%. The method does not work properly when the color of background is complex such as outdoor where the illumination changes drastically. Chih-Wen Su et al [3] studied the detection of people using the symmetry property of the human torso. The process can solve the problem of the incomplete extraction of the human silhouette. The accuracy of the result is 85%. Most of the missed detections were caused by the vertical occlusion. In this case, people who were occluded by the others only showed head and shoulders in the video frame. Jae-Won Kim et al [4] count the number of moving people entering the security door, which are detected and tracked using a fixed single camera. The object is detected using background subtraction, and then tracked. The bounding box allows partial object occlusion reducing the false warning. Tarek Yahiaoui et al [5] count the passengers in buses based on stereovision, segmentation, tracking and counting from original stereoscopic images, a dense disparity map which is converted into a height map. The height maps are segmented in order to highlight the passenger's heads at different levels. The Kalman filtering is applied for tracking, and followed by the counting step. Our research is inspired form the work of Tarek that performs good results.

3. Proposed System

In this section we introduce our proposed system for people-counting from the top view using depth and color images. Firstly, the depth images acquired from Kinect camera are analyzed for detecting moving objects using the background subtraction technique. Then, the heads of person are identified by object segmentation in the U-disparity representation. Next, the corresponding color

ROI of the heads is located from the color image of Kinect via camera calibration. The depth-color feature of the head is then tracked using the particle filtering method. And finally, the head tracking is used for the counting process. The overview of the proposed system is shown in the figure 1.

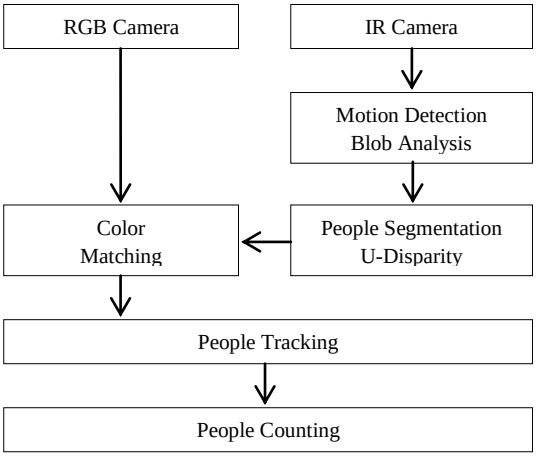


Figure 1. Proposed System

3.1 Moving Object Detection

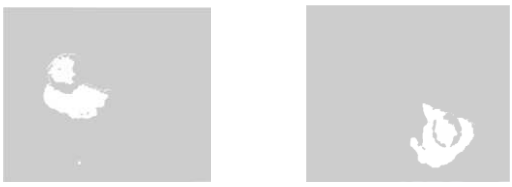
In our system, the motion detection is used to segment the moving object in depth image using background subtraction technique. In our particular problem, the motion of people may occur in the full area of the observation view. By performing motion detection only on color image, we found that the result is inapplicable. In contrast, motion detection on depth image is more reliable than on color image, noted that the change of the illumination or motion has strong impact on the intensity of color image but not on the depth information. In our system, we apply the Gaussian Mixture Model (GMM) method [6] for background modeling. The background is modeled by computing its probability $P(X_t)$ as the following:

$$P(x_t) = \sum_{i=1}^{K} \omega_{i,t} * \eta(x_t, \mu_{i,t}, \sum_{i,t}) \qquad (1)$$

where $\omega_{i,t}$ is weighted Gaussian, K is number of Gaussian (fixed to 5), And η is probability of Gaussian function defined by:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{\frac{-1}{2}} (X_t - \mu_t)^T \Sigma^{-1} (X_t - \mu_t) \qquad (2)$$

The result of background subtraction is shown in figure 2. We can notice that motion object can be well segmented in the most cases that separate moving person from the floor.



(a) (b)

Figure 2.the result of motion detection

After motion detection, the blob analysis is applied in order to find the bounding box of the moving person. This bounding box contains certainly peoples, but the number is unknown, which makes necessary a segmentation step.

3.2 People Segmentation

In this step, each blob will be studied in detail in order to compute the exact number of people with its regions; one blob may be split into several blobs. To achieve this goal, we use the U-disparity representation as an intermediate transformed domain.

U-disparity is estimated from disparity map on x-axe defined by the following equation:

$$V(I_\Delta) = I_{\Delta u} \qquad (3)$$

where $I_{\Delta u}$ is u-disparity, and V is a vertical accumulate operator of similar intensity on each column of disparity map. Figure 3 shows results of u-disparity for the moving object of one blob. From the results, we can notice that the u-disparity can discriminate the upper structure of human in the occlusion motions by its depth information from the top-view.

(a) (b)

Figure 3. the U-Disparity

In each blob, by using this u-disparity the head positions are marked by selecting the local maximums of depth values. Then, their sizes are considered in order to verify the real number of person. And the blob is split into the corresponding number.

Finally, the color from color image is then matched to each blob. Colors can represent the parts of body depending on hair or cloth colors, which is necessary for the tracking process. This step is done by segmenting the moving person in color image that correspond to the person found in the depth image via image transformation, using camera calibration matrix. Firstly, the depth pixels will be transformed into 3D points P3D.x, P3D.y, P3D.z using the following equations:

$$P3D.x = (x_d - cx_d) \times \frac{(depth(x_d, y_d))}{fx_d} \qquad (4)$$

$$P3D.y = (y_d - cy_d) \times \frac{(depth(x_d, y_d))}{fy_d} \qquad (5)$$

$$P3D.z = depth(x_d, y_d) \tag{6}$$

Where fx_d, fy_d, cx_d, cy_d are intrinsic of IR camera

fx_d =5.9421434211923247e+02

fy_d =5.9104053696870778e+02

cx_d =3.3930780975300314e+02

cy_d =2.4273913761751615e+02

Then, the 3D points are converted to 2D points in color image.

$$\begin{bmatrix} P3D'.x \\ P3D'.y \\ P3D'.z \end{bmatrix} = R.\begin{bmatrix} P3D.x \\ P3D.y \\ P3D.z \end{bmatrix} + T \tag{7}$$

$$P2D_{rgb}.x = \frac{P3D'.x * fx_{rgb}}{P3D'.z} + CX_{rgb} \tag{8}$$

$$P2D_{rgb}.y = \frac{P3D'.y * fy_{rgb}}{P3D'.z} + CY_{rgb} \tag{9}$$

where R and T are the rotation and translation matrix respectively.

$$R = \begin{bmatrix} 9.9984628826577e-01 & 1.2635359098409e-03 & -1.7487233004433e-02 \\ -1.4779096108364e-03 & 9.9992385683545e-01 & -1.2251380107635e-02 \\ 1.7470421412464e-02 & 1.2275341476520e-02 & 9.9977202419948e-01 \end{bmatrix}$$

$$T = \begin{bmatrix} 1.9985242312092e-02 \\ -7.4423738761613e-04 \\ -1.0916736334322e-02 \end{bmatrix}$$

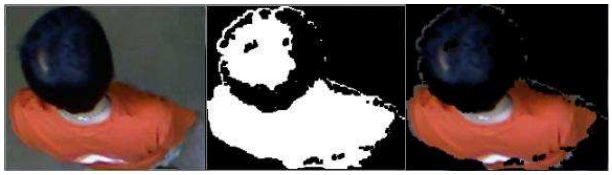
fx_{rgb} =5.2921508098293293e+02

fy_{rgb} =5.2556393630057437e+02

cx_{rgb} =3.2894272028759258e+02

cy_{rgb} =2.6748068171871557e+02

Finally, the transformed depth pixels, figure 4(b), and color pixels, figure 4(a), is matched with a simple *and* operator, which produces the result shown in figure 4(c).



(a) (b) (c)

Figure 4. (a) color image (b) depth image (c) color object

3.3 People Tracking

From the last section, in each frame, the moving persons are segmented as blobs with it properties, such as, centroid, and color pattern. In our tracking method, we combine two techniques for a better real time tracking. Firstly, the Kalman filter is applied for predicting the movement of the person in frame i+1 using centroids of the previous frames. Secondly, the nearest blob to the predicted one of the frame i+1 is selected, and then compared to the current blob of the frame i with its similarity using normalized color pattern correlation. If its similarity is acceptable comparing to a threshold, we believe that the selected blob is the right tracking person. If not, the other blob will be selected ordering it by its distance to the predicted blob. The process will be repeated until finding the optimal one. The linear stochastic differential equation of Kalman filter is defined by the following:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{10}$$

$$\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k,i} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k-1,i} + w_{k-1,i} \tag{11}$$

Where x_k (x, y, v_x, v_y) are the centroid and the velocity of blob in the kth frame respectively. Here, we assume that the motion of blob between two successive frames can be uniformly approximated as straight line; the frame interval Δt is very short. For the measurement, z_k is defined as the detected blob in frame $(k+1)^{th}$ that is

$$z_k = Hx_k + v_k \tag{12}$$

The random variables w_k and v_k represent Gaussian noise of the process and measurement respectively. The algorithm of Kalman filter estimates a process by using feedback control technique: estimating the process state by an appropriate model and doing feedback by noisy measurements. As such, the equations of Kalman filter are formed into two groups: prediction and correction equations. In the post tracking of hand, the algorithm can be described as the following figure 5.

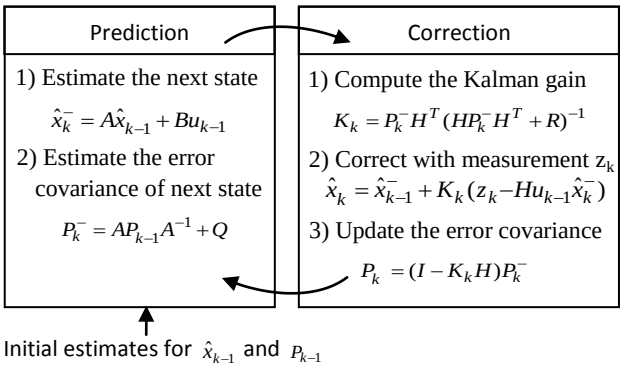


Figure 5. The operation of Kalman filter.

3.4 People Counting

In this section, we explain how to count the tracking people. The figure 6 shows the counting state machine. Firstly, all blobs in the field of view will be tracked using the method explained above. Each blob will have three statuses: alive, counted, and dead. Alive status is distributed to the blob when its presence on the scene is greater than a threshold *th1*. Counted status is defined for the blob when its presence on the counting area is greater than a threshold *th2*. And the dead blob is when people go out from the field of view.

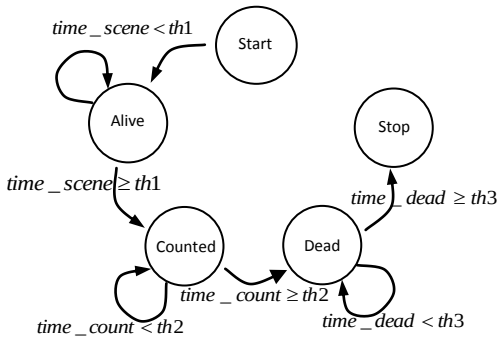


Figure 6. Counting finite state machine.

4. Experimental results

The dataset is acquired from Kinect camera installed on the top of the main entrance of the Faculty of Engineering, Prince of Songkla University. The color and depth image is analyzed at the same time via RGB camera and IR camera, shown in figures 7. The dataset is done using our adaptive version of CL NUI Platform with resolution 320*240 pixels. People walk through the counting area, setup horizontally at the middle of the field of view. Four groups (1 person, 2 persons, 3 persons, 4 persons respectively) are testing with two directions (in and out). Testing system is run on Intel processor Dual Core, 2.4 GHz, 1 GB memory.

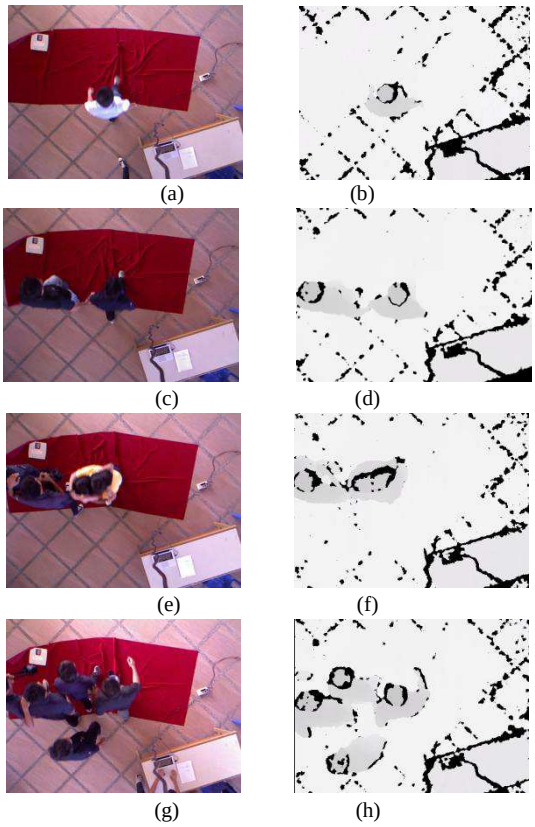


Figure 7. Testing dataset (a) (b) 1 person, (c) (d) 2 persons, (e)(f) 3 persons, (g)(h) 4 persons.

In our experimentation, the table 1 shows the correction rate with respect to the number of people in the testing cases.

Groups	Time	Persons	Detected	Correct rate
1 person	20	20	20	100.00
2 persons	10	20	19	95.00
3 persons	6	18	16	88.89
4 persons	4	16	14	87.50
Cases	**Time**	**Persons**	**Detected**	**Correct rate**
In	20	37	35	94.59
Out	20	37	34	91.89
Average	**40**	**74**	**69**	**93.24**

Table1: the correction rate of our counting system

We can note that globally our system can perform very good correction rate in most of the cases at around 93%. However, the system cannot count correctly when the heads of two persons are connected physically. Additionally, when two persons have mostly the same height, our system cannot well identify them.

5. Conclusion

In this paper, we presented a people counting system from the top-view images using depth and color information. The U-disparity is specially applied for segmenting the occluded people, which is tracked by a mixture of position and color properties. The system provides a good recognition rate at 93.0%. In the future, some additional techniques for head segmentation could be applied for increasing the accuracy of the system.

References

[1] Antonio Albiol, Inmaculada Mora, and Valery Naranjo, “Real-Time High Density People Counter Using Morphological Tools”, IEEE Transactions on Intelligent Transportation Systems, Barcelona, VOL. 2, NO. 4, pp. 204-218, Dec 2001.

[2] Xi Zhao, Emmanuel Dellandrea and Liming Chen, “A People Counting System based on Face Detection and Tracking in a Video” , Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, Genova, pp.67-72, 2-4 Sept 2009.

[3] Chih-Wen Su, Hong-Yuan Mark Liao and Hsiao-Rong Tyan, “A Vision-based People Counting Approach Based on the Symmetry Measure”, IEEE International Symposium on Circuits and Systems, Taipei, pp. 2617-2620, 24-27 May 2009.

[4] Jae-Won Kim, Kang-Sun Choi, Byeong-Doo Choi, and Sung-Jea Ko, “Real-time Vision-based People Counting System for the Security Door”, ITC-CSCC :International Technical Conference on Circuits Systems, Computers and Communications, Korea, pp. 1418-1421, 2002

[5] Tarek Yahiaoui, Cyril Meurie, Louahdi Khoudour, and Francois Cabestaing , “A People Counting System Based on Dense and Close Stereovision”, Image and Signal Processing: 3rd International Conference, ICISP 2008, France, pp.59-66, 2008

An Image Segmentation Method for Glaucoma Detection using the CDR

Wanicha Ruengkitpinyo† **Toshiaki Kondo**† **Waree Kongprawechnon**†
†School of Information, Computer, and Communication Technology (ICT),
Sirindhorn International Institute of Technology (SIIT),
Thammasat University, Bangkok, THAILAND
{ru.wanicha@gmail.com, tkondo@siit.tu.ac.th, waree@siit.tu.ac.th}

Abstract

Glaucoma is a serious eye disease, overtime it will result in gradual blindness. Early detection of the disease will help prevent against developing a more serious condition. Cup-to-Disc ratio or CDR is an important clinical indicator for glaucoma diagnosis. Nowadays, CDR is manually determined by trained ophthalmologists and has to be diagnosed by using retinal images from special cameras which are expensive. These are limitations of potential in mass screening for early detection. This paper focused on developing system that can provide suitable way to compute CDR result which give the most accurate result and prevent from current limitation. The edge detection with power-law transformation method is described for both optic cup and optic disc segmentation.

Keywords: Cup-to-Disc, Optic cup, Optic Disc, Edge Detection, Power Law

1 Introduction

Glaucoma is a disease of increased pressure within the eyeball. The increase in pressure result in immoderate amount of stress to be put to the attachment of the optic nerve to the eye [1] [2]. Lack of treatment for glaucoma can lead to permanent blindness, which glaucoma is the second leading cause worldwide. Early detection of the disease will help prevent against developing a more serious condition.

For CDR, diameter of both optic disc and optic cup are used. Optic disc is usually known as blind spot. There is no cell to detect light on optic disc. Optic disc is bright area which locates at the back of eyeball, while optic cup is the brightest area within an eyeball. Optic cup diameter for normal case and glaucoma case are shown in Figure 2 [3]. The cup diameter is larger, corresponding to the loss of vision. Vertical cup-to-disc ratio value that is greater than or equal to 0.65 is considered to be suspicious for glaucoma [4] [5].

Figure 1. State of human vision.

1.1 Problem Statement

Optical Coherence Tomography or OCT is widely used for medical imaging techniques. OCT is an optical signal acquisition and processing method which captures optic images in three-dimensional image. Both OCT and ophthalmologists are having the same problem which are insufficient and expensive. They can only find in large hospitals or private hospitals. Some of glaucoma patients are insufficient in funds. This can make patients get a late glaucoma treatment and bad for their eye health. The need of experts and large amount of expenses are limitations of potential in mass screening for early detection.

To solve problem of OCT images, fundus images are selected to be input images for this project. Fundus camera is another type of camera which can use to capture retina images. Fundus camera has less expensive costs compar-

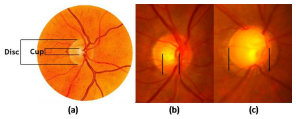


Figure 2. Components of eyeball (a) Optic cup and optic disc area within eyeball. (b) Normal optic nerve cell with small optic cup. (c) Glaucomatous optic nerve head.

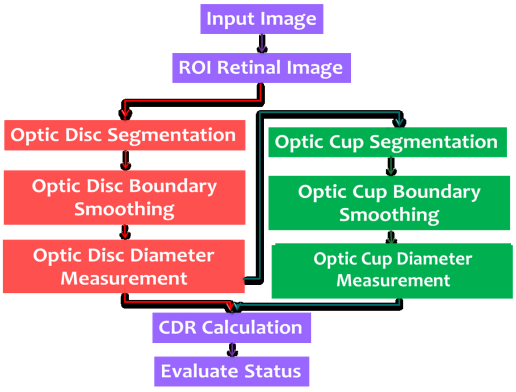


Figure 3. Process of this paper.

ing with OCT images. It can be find in every province medical center, not only large hospitals as same as OCT. Fundus image can be used for glaucoma diagnosis via CDR method.

1.2 Objectives

The purpose of this paper is to developing system that can provide suitable way to compute CDR results which give the most accurate result. Accuracy of results is the most important indicator for early glaucoma detection which can prevent against developing to more serious condition.

2 Methodology

2.1 System Configuration

For CDR calculation, both optic cup and optic disc diameter have to be measured. Figure 3 demonstrates process of this paper.

2.2 Region of Interest Detection

To calculate the vertical cup to disc ratio (CDR), the optic cup and disc first have to be segmented from the retinal images. Input retinal images which are used in this paper are taken by Fundus camera. The size of original retinal images are very large, 3072 x 2048 high-resolution digital image, as shown in Figure 4 (a). Computing time for directly apply proposed method to original retinal images is also large. Therefore, size reduction of original retinal images is needed.

In this paper, the interested region is only at the location of optic disc and optic cup. By using ROI images, it takes less computing time because it has smaller size [6]. According to our experiment, optic cup and optic disc are in top

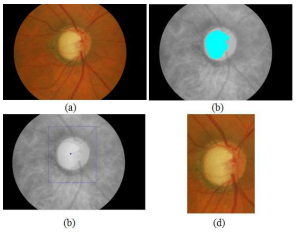


Figure 4. Process of cropping region of interest. (a) Original retinal image of size 3027x2048 pixels (b) Top two third intensity detected (blue area) (c) ROI localization (d) The ROI image of size 480x750 pixels

two-third intensity in retinal image. Localizing top two-third intensity in retinal image[6] (Figure 4 (b)) is the first step for crop only region of interested. Next, localizing region of interested, as shown in Figure 4 (c), by combine region in Figure 4 (b) with its surrounding pixels. Finally, the ROI image is returned as an image of size 480x750 pixels, Figure 4 (d).

2.3 Optic Disc Segmentation

ROI image will be read in red channel because optic disc region will be clearly seen in red channel, according to our experiment. In this step, optic disc diameter will be measured.

Power-law transformation, Equation (1), is used to solve brightness problems of retinal images. If $0 \leq \gamma \leq 1$ is selected, the image will be brighter. In another hand if $\gamma < 1$, result image will be vice versa. In this paper, γ = 5 is selected. The pixel value vs maximum intensity of pixel graph is plotted after power law is applied to the retinal image. The location of optic disc will be shown in Figure 6. Retinal image after power law transformation is applied and intensity graph are demonstrated in Figure 5.

$$s = Cr^{\gamma} \tag{1}$$

where s is output value, r is input value, C and γ are constant.

Next step, retinal image will be converting into black and white image. The reason is because intensities in retinal image are not strongly contrasts. Converting it into black and white image first can be solved this problem. Then canny edge detection is applied, where the threshold value for canny edge detection will be selected from Figure 5 (b).

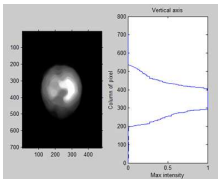


Figure 5. Power Law Transformation for Disc segmentation (a)Image after applied power law transformation (b) Graph of pixel value vs maximum intensity of pixel

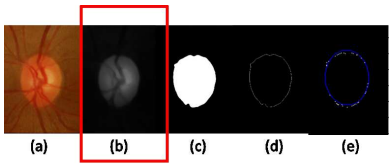


Figure 6. Step of Edge Detection with Power Law Transformation Approach for Disc segmentation (a) Input retinal image (b) Retinal image after power law transformation is applied (c) Black and white image (d) Edge image (e) Edge image with boundary

Output image with edge boundaries will be shown on the screen as shown in Figure 6 (d). Then the application of direct ellipse fitting is used to smoothing edge boundaries and optic disc diameter will be measured. The output is shown in Figure 6 (e).

2.4 Optic Cup Segmentation

For our experiment, optic cup is clearly seen in green channel. Therefore, ROI image will be read in green channel. Diameter of optic cup will be measured in this step.

All of the processes for optic disc segmentation will be also applied to optic cup segmentation. Retinal image after power law transformation is applied is shown in Figure 7. Figure 8 demonstrates all step of optic cup detecting by using our proposed approach.

3 Results

44 input retinal images from Mettapracharak hospital are used to evaluate the performance. 29 of the retinal images are normal patients with no clinical signs of glaucoma and 15 of retinal images are from patients with glaucoma. CDR values from Mettapracharak Eye Center are set as reference to find percentage error of CDR value from proposed approach.

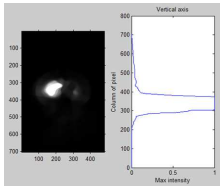


Figure 7. Power Law Transformation for Cup segmentation (a)Image after applied power law transformation (b) Graph of pixel value vs maximum intensity of pixel

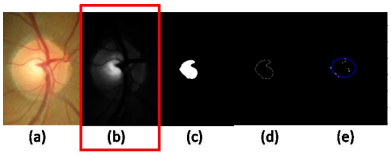


Figure 8. Step of Edge Detection with Power Law Transformation Approach for Cup segmentation (a) Input retinal image (b) Retinal image after power law transformation is applied (c) Black and white image (d) Edge image (e) Edge image with boundary

3.1 Optic Disc and Optic Cup Segmentation

The 44 images are processed using edge detection with power law transformation approach to obtain the optic disc and optic cup diameter for each retinal image. Three sample retinal images of optic disc are shown in (Figure 9) and (Figure 10) is for optic cup. The first row represents clinical result and second row represents the proposed approach result. The average error of the proposed method to optic disc and optic cup detection are 2.49% and 5.8%, respectively.

3.2 CDR Calculation

The distribution of CDR values from proposed method compares with clinical CDR values are shown in Figure 11. The average CDR error of

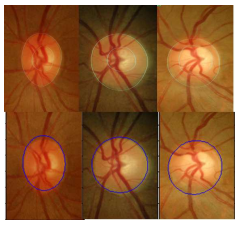


Figure 9. Optic disc segmentation (First row) Clinical optic disc (Second row) Edge detection with power law transformation optic disc

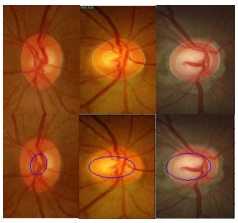


Figure 10. Optic Cup segmentation (First row) Clinical optic cup (Second row) Edge detection with power law transformation optic cup

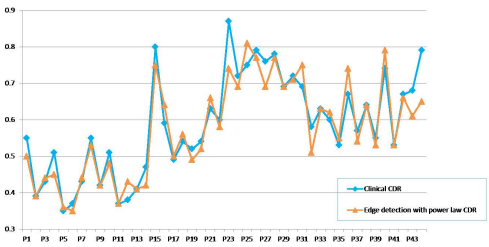


Figure 11. Compairson of CDR value between clinical result and edge detection with power raw transformation approach.

proposed method is 5.14%.

3.3 Glaucoma Diagnosis

The value of CDR which is more than 0.65 is used to assess a patient as a possible glaucoma case. To measure the performance of proposed method confusion matrix is introduced. A confusion matrix contains information about actual and predicted classifications.

There are others three variable which also used to evaluate proposed method results; Sensitivity, Specificity and Accuracy. Accuracy is the percent correctness of all samples tested. Sensitivity is the percent correctness of true sample, which are glaucoma cases in this paper. Specificity is the percent correctness of the false, which are normal cases. Figure 12 shows the confusion matrix of glaucoma classification for this paper.

Purposed Method Results		Clinical Results: Normal	Clinical Results: Glaucoma
	Normal	28	1
	Glaucoma	1	14

Specificity: 97%
Sensitivity: 87%
Accuracy: 95%

Figure 12. confusion matrix of glaucoma classification.

4 Conclusion

The CDR is an important clinical indicator for determine the risk of the presence of glaucoma in an individual. In this paper, we have presented a method to calculate the CDR automatically from fundus images. The performance of our approach is evaluated using proximity of the calculated CDR to the manually graded CDR. The percentage error by using our proposed method for optic disc segmentation, optic cup segmentation and CDR are 2.49% , 5.8% and 5.14%, respectively.

Acknowledgment

This research is financially supported by Center of Excellence in Biomedical Engineering, Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU) and the National Research University Project, Thailand Office of Higher Education Commission.

References

[1] Johns Hopkins Medicine. Glaucoma. `http://www.hopkinsmedicine.org`, October 2012. Accessed on 29-December-2012.

[2] North Cinicinnati Eyecare. `http://www.northcincinnatieyecare.com`. Accessed on 30-January-2013.

[3] MedFlux: Medical Photo Gallery. `http://www.medflux.com`. Accessed on 29-December-2012.

[4] D.Wong J. Liu, J. Lim, X. Jia, H. Li F. Yin, and T.Wong. Optic cup and disk extraction from retinal fundus images for determination of cup-to-disc ratio. *Industrial Electronics and Applications*, June 2008.

[5] Glaucoma Clinical Practice Guideline. `http://www.moh.gov.sg`, October 2005. Accessed on 30-December-2012.

[6] Beng Hai Lee, Jiang Liu, Wong, D.W.K., Ngan Meng Tan, Joo Hwee Lim, Fengshou Yin, Weimin Huang, Huiqi Li, and Tien Yin Wong. Optic disc region of interest localization in fundus image for glaucoma detection in argali. *Industrial Electronics and Applications*, June 2010.

Automatic Initialization of Multiple Active Contours Applied to Ultrasound Images of Breast Cancer

Khwunta Kirimasthong†* Stanislav S. Makhanov† Annupan Rodtook‡

†School of Information, Computer and Communication Technology,
Sirindhorn International Institute of Technology, Thammasart University, Thailand
‡Department of Computer Science, Ramkhamhaeng University, Thailand
khwunta.kir@mfu.ac.th

Abstract

The multiple active contours (snakes) are widely used in image processing to extract objects of interest. However, initialization of the active contours is still an open problem. This paper proposes a new completely automatic initialization of snakes based on a combination of Phrase Portrait Analysis (PPA) image intensity clustering and multi-stage evolution of the active contours. The proposed method has been tested with the ultrasound images of breast. The experiments show that the proposed automatic initialization is applicable to noisy ultrasound images.

Keywords: Automatic initialization, Active contours, Snakes, Phrase portrait analysis, Generalized gradient vector flow

1 Introduction

The accuracy of the computer based diagnostics of ultrasound (US) images is still not sufficient. It is often difficult to separate the tumor region from the background even when the existence of the tumor is evident.

Among the most promising techniques for extraction of complex objects from digital images are active contours or snakes, originally introduced by Kass et al. [1], [2]. However, the success of the snake based segmentation depends strongly on initialization [3-4] that is on the position where the snake is initially placed.

Up to now several solutions of the problem have been proposed and implemented. A quasi-automatic initialization that relies on the divergence of gradient vector flow field is proposed in [5]. The mean shape algorithm for automatic initialization is used in [6]. The Poisson inverse gradient for automatic active contour initialization is introduced in [7]. Finally, a problem dependent solution based on self-organizing maps [8] has been proposed for extraction roads in satellite images.

However, the proposed methods do not show display proper results when applied to US images of the breast cancer. Therefore, the main contribution of this paper is a completely automatic initialization model based on the combination of PPA, image intensity clustering and active contours applied using a multi-stage evolution algorithm.

2 Methodology

Our proposed algorithm consists of five steps. First, the PPA [3] is combined with edge-based detection in the preprocessing sequence that includes a speckle noise reduction, median filtering, contrast enhancing and region growing. Next the general gradient vector flow (GGVF) [3] is applied to the resulting image (Figure 1). Furthermore, we evaluate the PPA map which is clustered and classified as the background, boundary and the noise. Next, we perform a geometric analysis of the pixels with a strong PPA score and set up seeds for and multiple quadratic active contours [4]. Finally, we run the active contours using a special multi-stage evolution algorithm.

Consider now the geometric analysis illustrated in Figure 1. The PPA map clustered into the boundary points, noise points and background points. Next, we using the assumption that the tumor is in general round we construct a convex hull based on the candidate boundary points. Note that it is possible that some candidate boundary points belong to the false boundaries represented by speckle noise and shadows. However, our simple procedure makes it possible to confine the tumor within the convex hull

(Figure 1 (b)). Next, we select the noise points inside the convex hull positioned far from the boundary. These points are the seeds for the quadratic multiple snakes. In other words the snakes have been initialized at the points of convergence and noise inside the tumor (Figure 3 (a)). This approach makes it possible to avoid false boundaries inside the tumor since the false boundaries are already inside the initial snakes (Figure 3(b)).

We run the snakes until convergence using the anti-gradient. This is because the convergence and noise points are represented by attracting (not repelling) stars. At this stage the snakes converge and stop at the false boundaries between the attracting stars and the background. Therefore, we offset the snakes and use the original gradient vector field to evolve the snake to the true boundary. The algorithm is illustrated in Table 1.

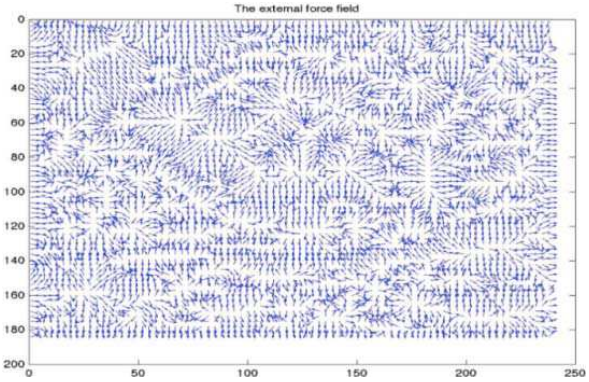
3 Experiments and Results

The algorithm has been tested by two synthetic tumors and three real tumors. The automatic initialization works well and the desired boundary was successfully extracted in every case. The PPA score corresponding to the boundary was 0.9.

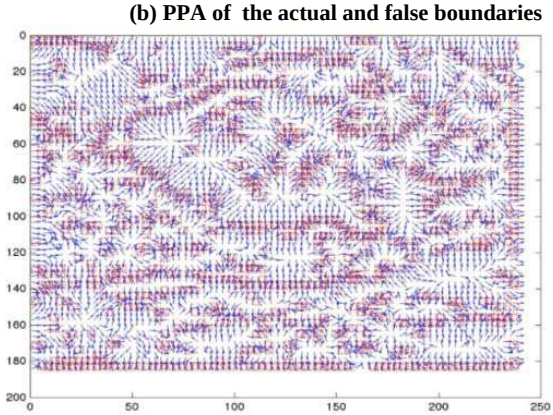
4 Conclusion

The proposed initialization model based on a combination of the PPA and the image intensity classification works well to extract the boundary of breast cancer tumor in both synthetic and real ultrasound images. The results make it possible to conjecture that the proposed techniques will further succeed in segmentation of a variety of tumors displayed in ultrasound images of the breast.

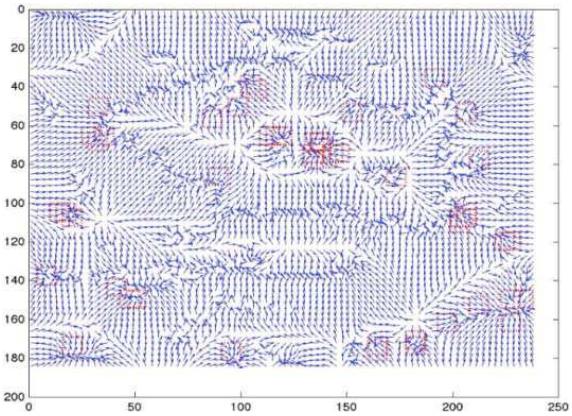
(a) GGVF of the original image



(b) PPA of the actual and false boundaries



(c) PPA of noise points



(d) Final GGVF

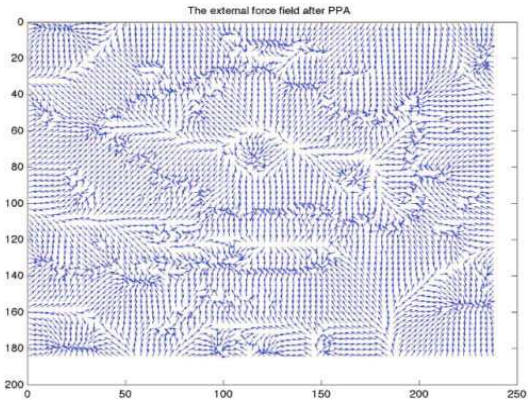
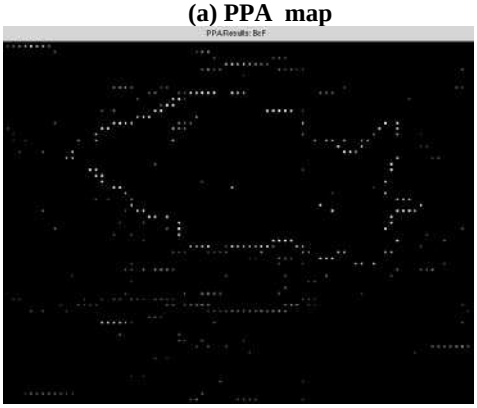
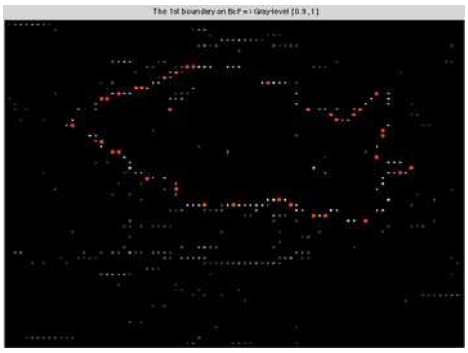


Figure 1. The GGVF fields and the PPA analysis

(a) PPA map



(b) PPA strong boundary points



(c) PPA weak boundary and noise

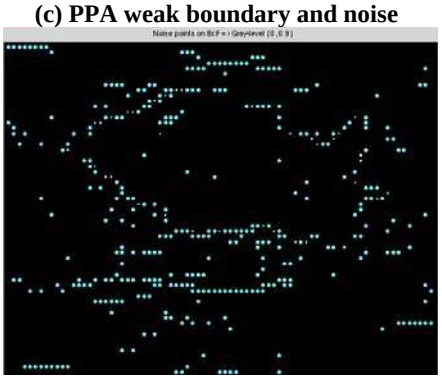
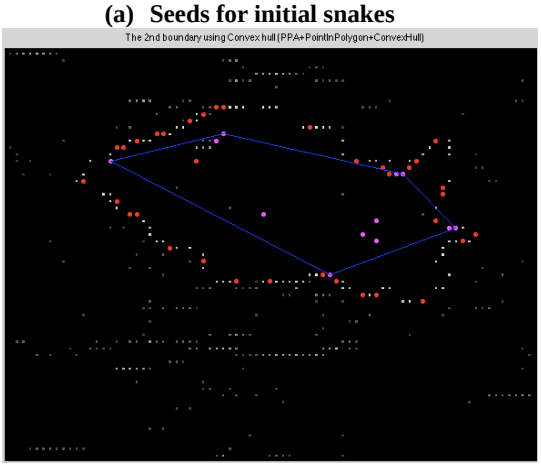


Figure 2. PPA classification based on the image intensities

(a) Seeds for initial snakes



(b) Initialized snake

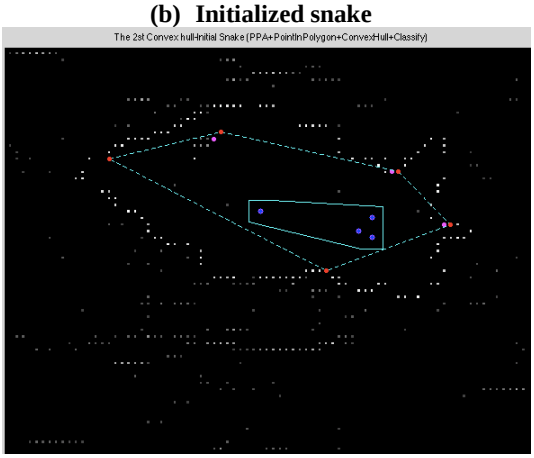
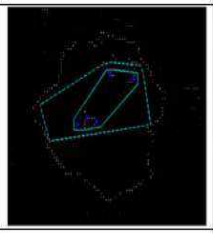
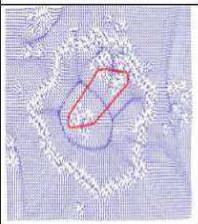
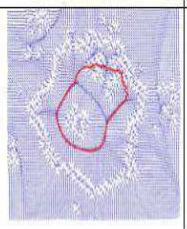
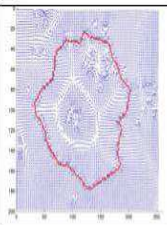
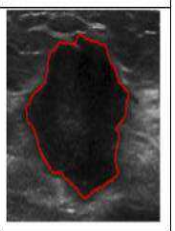
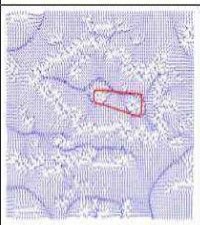
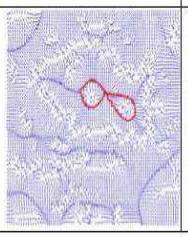
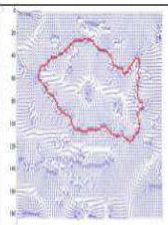
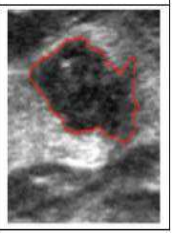


Figure 3. The automatic initialization based on PPA

Table1. Extraction results

Image	Point in Polygon Checking	The 1st Initial snake	The 1st round snake extraction	The 2nd round snake extraction	The result on input image
T2org.bmp					
					
T3org.bmp					
					

References

[1] M. Kass et al. Snakes: Active Contour Models, *International Journal of Computer Vision*, 1(4):321-331, 1988.

[2] C. Xu and J. L. Prince. Gradient vector flow: a new external force for snakes, *IEEE Conference on Computer Vision and Pattern Recognition*, 66–71, 1997.

[3] S. Chucherd and S.S. Makhanov. Sparse Phase Portrait Analysis for Preprocessing and Segmentation of Ultrasound Image of Breast Cancer, *IAENG International Journal of Computer Science*, 38(2): 146-159, 2011.

[4] Y. Nakaguro, S. S. Makhanov and M.N. Dailey. Numerical Experiments with Cooperating Multiple Quadratic Snakes for Road Extraction, *International Journal of Geographical Information Science*, 25(5): 765-783, 2011.

[5] C. Tauber et al. Quasi-automatic initialization for parametric active contours, *Pattern Recognition Letters*, 31, 83-90, 2010.

[6] S. Choi and C. Kim. Automatic Initialization Active Contour Model for the Segmentation of the Chest Wall on Chest CT, *Healthcare Informatics Research*, 16(1):36-45, 2010.

[7] B. Li and S.T. Acton, Automatic Active Model Initialization via Poisson Inverse Gradient, *IEEE Transactions on Image Processing*, 17(8):1406-1420, 2010.

[8] R.G. Lotte et al. Road Centre-Axis extraction in Airborne SAR images: An Approach Based on Active Contour Model with the Use of Semi-automatic Seeding, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1/W1:207-212, 2013.

Multiple Quadratic Snakes for Segmenting Tumors in Medical Images

Yoichi Nakaguro[†] **Stanislav Makhanov**[†]
[†]Sirindhorn International Institute of Technology
Thammasat University, Thailand
{ynakaguro,makhanov}@siit.tu.ac.th

Abstract

Classical active contour models provide an elegant framework for optimal estimation in image processing. However, their adaptability to image noise and convergence speed have been major challenges. In this paper, we address the problems by combining advanced active contour techniques such as Gradient Vector Flow (GVF), quadratic snake interaction, and the adaptive initialization of multiple snakes. Through two independent experiments, we show that our snake model can effectively segment objects in images with high level noise such as CAPTCHAs and medical images. First, we show that our algorithm is fast enough for practical applications by measuring the time efficiency in successful segmentation experiments of a CAPTCHA challenge data set. Second, we consider medical image segmentation that requires high degree of noise tolerance. We test our algorithm against a tumor image with higher noise than CAPTCHAs. With an adaptive initialization approach, our preliminary segmentation results shows our technique is effective to the application to real medical images with tumors.

Keywords: Active Contours, Quadratic Snakes, Snake Multiplicity, GVF, Adaptive Initialization

1 Introduction

Flexible automated systems capable of extracting structures and regions of interest from digital images encompass a broad range of applications. For example, applications such as automated road extraction and object detection in medical images would boost the productivity of technicians enormously. Manual marking and extraction of those objects is an extremely slow and laborious process considering the degree of complexity that those objects can possess in the real world. Towards the ultimate goal of fully automated image object extraction, there has been a great deal of progress in the computer vision and image processing communities. But after 30 years of research, (see [1] for a thorough review), there is still no system attaining the speed, robustness, and level of automation necessary for practical application on arbitrary imagery.

Classical active contour models [2; 3; 4] provide an elegant framework for optimal estimation in image processing; rather than writing an algorithm to extract the object or region of interest, we simply consider an energy functional a minimum of which is achieved at a good solution. Then, given a new image, we use general optimization techniques to find a contour minimizing the energy functional.

Although the classical models should fundamentally work well, in practice, their adaptability to image noise and convergence speed become major challenges. For noise, image enhancement based on object types such as oriented filtering applied to road extraction applications, can greatly improve extraction results [5]. For convergence speed, Xu and Prince's gradient vector flow (GVF) [6] technique makes it possible to achieve fast convergence of snakes initialized far from the boundary of the object. GVF also improves convergence to objects with concave boundaries where the classical snake models fail.

Rochery et al. have proposed a geometric model for higher-order active contours, in particular quadratic snakes, for extraction of linear structures like roads [7]. The idea is to use a quadratic formulation of the contour's geometric energy to encourage anti-parallel tangents on opposite sides of a road and parallel tangents along the same side of a road. These priors increase the final contour's robustness to partial occlusions and decrease the likelihood of false detections in regions not shaped like roads.

In this paper, we propose a multiple quadratic multiple snake model that uses the GVF external force. We show that our approach is fast enough for practical applications by measuring the time

efficiency of our snake model in successful segmentation experiments of a CAPTCHA challenge data set. We also show preliminary segmentation results in the application for real medical images with tumors using our adaptive initialization technique.

2 Methods

2.1 Quadratic snake model

This section briefly overviews our quadratic snake model by extending the original snake model [2] with quadratic geometric energy [7].

We define a closed 2D spline $\vec{\gamma}(p) = [x(p), y(p)]^T, p \in [0,1]$ that minimizes energy functional

$$E(\vec{\gamma}) = \int E_s(\vec{\gamma}, \vec{\gamma}_p, \vec{\gamma}_{pp}, p)dp = E_g + E_i, \quad (1)$$

where $\vec{\gamma}_p$ and $\vec{\gamma}_{pp}$ are first and second partial derivatives of a contour $\vec{\gamma}$, respectively.

E_g is the geometric energy defined by

$$E_g(\vec{\gamma}) = \int \left(\frac{1}{2}\alpha\vec{\gamma}_p^2 + \frac{1}{2}\beta\vec{\gamma}_{pp}^2\right) dp - \frac{\delta}{2}\iint \vec{t}(p_1)\cdot\vec{t}(p_2)\ \Psi\left(\|\vec{\gamma}(p_1) - \vec{\gamma}(p_2)\|\right)\ dp_1\ dp_2, \quad (2)$$

The first integral represents the stretch energy and bending energy, respectively. α and β are the weights of those energy terms. A large stretch energy encourages the contour to be shortened, on the other hand, a large bending energy forces the contour to be straightened. $\vec{t}(p)$ is the unit-length tangent to $\vec{\gamma}$ at point $\gamma(p)$, and $\Psi(z)$, given the distance z between p_1 and p_2, is used to weight the interaction between those two points. For positive δ, $E_g(\vec{\gamma})$ is minimized by contours with short length and parallel tangents.

The quadratic term is responsible for interactions between points on the snake. The sigmoid function $\Psi(\cdot)$ defines the radius of the region in which anti-parallel tangents should be discouraged:

$$\Psi(z) = \begin{cases} 1 & \text{if } z < d - \epsilon, \\ 0 & \text{if } z > d + \epsilon, \\ \frac{1}{2}\left(1 - \frac{z-d}{\epsilon} - \frac{1}{\pi}\sin\pi\frac{z-d}{\epsilon}\right) & \text{otherwise.} \end{cases} \quad (3)$$

d is approximate width of the region of interest and ϵ is the sharpness of the sigmoid function.

E_i is the image energy that depends on the image intensity $I(x,y)$ as follows

$$E_i(\vec{\gamma}, I) = -\int \lambda|\nabla I|^2 dp, \quad (4)$$

where λ is a weight.

Contour γ is a solution of the following Euler equation:

$$-\frac{d^2}{dp^2}\left(\frac{\partial E_s}{\partial\vec{\gamma}_{pp}}\right) + \frac{d}{dp}\left(\frac{\partial E_s}{\partial\vec{\gamma}_p}\right) - \frac{\partial E_s}{\partial\vec{\gamma}} = 0. \quad (5)$$

2.2 GVF external force

To improve the convergence, we adopt Xu and Prince's gradient vector field (GVF) [6] technique. The GVF is a vector field $\vec{V}(x,y) = \left[u(x,y)\ v(x,y)\right]^T$ that minimizes an energy functional given by

$$E(\vec{V}) = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla I|^2|V - \nabla I|^2\ dx\ dy, \quad (6)$$

where

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \quad v_x = \frac{\partial v}{\partial x}, \quad v_y = \frac{\partial v}{\partial y}.$$

Minimizing the functional leads to an Euler equation given by

$$\mu\nabla^2\vec{V} - (\vec{V} - \nabla I)|\nabla I| = 0. \quad (7)$$

The Laplacian creates a slow varying field in homogeneous regions causing elimination (diffusion) of noise and false boundaries. With large $|\nabla I|$, the energy is dominated by the second term. Therefore the energy achieves the minimum, when $\vec{V} = \nabla I$, encouraging the snake to snap to the object boundary.

2.3 Snake multiplicity

A single quadratic snake [7] is unable to extract enclosed regions and multiple disconnected characters in an image. We address this limitation by introducing a *family* of cooperating snakes that are able to split, generate offspring (anti-snakes), disappear, and merge if necessary.

In our formulation, specifying the points on $\vec{\gamma}$ in a counterclockwise direction creates a *shrinking snake*, whereas specifying the points on $\vec{\gamma}$ in a clockwise direction creates a *growing snake*.

2.3.1 Splitting a snake

We split a snake into two snakes whenever two of its arms are squeezed too close together.

2.3.2 Deleting a snake

A snake $\vec{\gamma}$ is deleted if it has perimeter less than L^{delete}.

2.3.3 Generating an anti-snake

The evolution of the snake continues until it either attaches itself to a boundary of the character or gets deleted. In the first case an anti-snake is generated by offsetting the original snake by d^{offset} pixels inside the object.

2.3.4 Merging two snakes

We merge two growing snakes when they come too close to each other.

2.4 Adaptive initialization

With the ability of anti-snake generation, it is possible that we simply initialize a single shrinking snake along the boundary of an image and extract all inner objects. For medical images with complex shape objects, however, more sophisticated initialization is generally required to obtain correct convergence. For this purpose, we employ an adaptive initialization approach in which we initialize multiple shrinking/growing snakes depending on the intensity of image regions.

3 Experiments

3.1 Time efficiency of multiple quadratic snakes

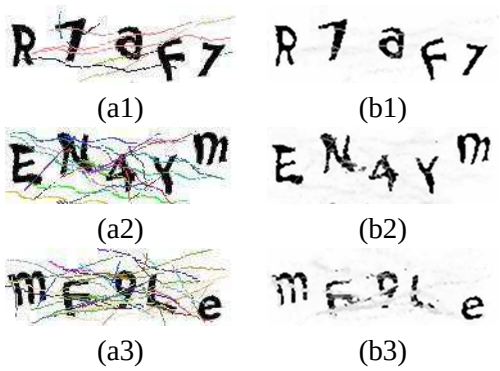


Figure 1. CAPTCHA samples. (a) Original CAPTCHAs. (b) Pre-processed CAPTCHAs.

In this experiment, we show that our approach is fast enough for practical applications by measuring the time efficiency of our snake model in successful segmentation experiments of a CAPTCHA challenge data set.

Table 1. Convergence time (sec) and accuracy vs noise levels

	Low	Medium	High
Ave	2.91	2.79	3.36
Std	0.31	0.23	0.51
Accuracy	0.8	0.9	0.4

To create the CAPTCHA image set used for the experiment, we used the popular Drupal CAPTCHA API [8], which produces CAPTCHAs characterized by distortion and tilting of the characters followed by the addition of line noise. We generated 480×160 CAPTCHAs each containing five characters. The average measured line noise levels are 6.5 dB, 5.2 dB, and 4.4 dB for Low, Medium, and High, respectively. After pre-processing the CAPTCHAs with an oriented filter followed by median filter, we measured the convergence speed and accuracy of our multiple snake algorithm against the CAPTCHA data set containing 10 samples in each noise level. The accuracy is determined by recognition rate using a SVM (Support Vector Machine) pre-trained by similar CAPTCHA data.

Table 1 presents results of the numerical experiment, that is, the pixel-wise extraction of CAPTCHAs corrupted by the three noise modes (Figure 1). We found that noise level Low and Medium show similar results (Medium is even slightly better than Low) due to the fact that the pre-processing results of both levels become quite comparable. In the context of CAPTCHA recognition by machines, even accuracy with 1% is considered to be a substantial threat to a system. Our proposed algorithm extracted CAPTCHA characters with practical accuracy and time cost.

3.2 Segmentation of a tumor in a medical image

In this experiment, we apply our multiple snake algorithm to segment a tumor from a real medical image. We tested two types of initialization approach. One is the simple initialization with a shrinking snake along the boundary of the image; and the other is manually setting multi-

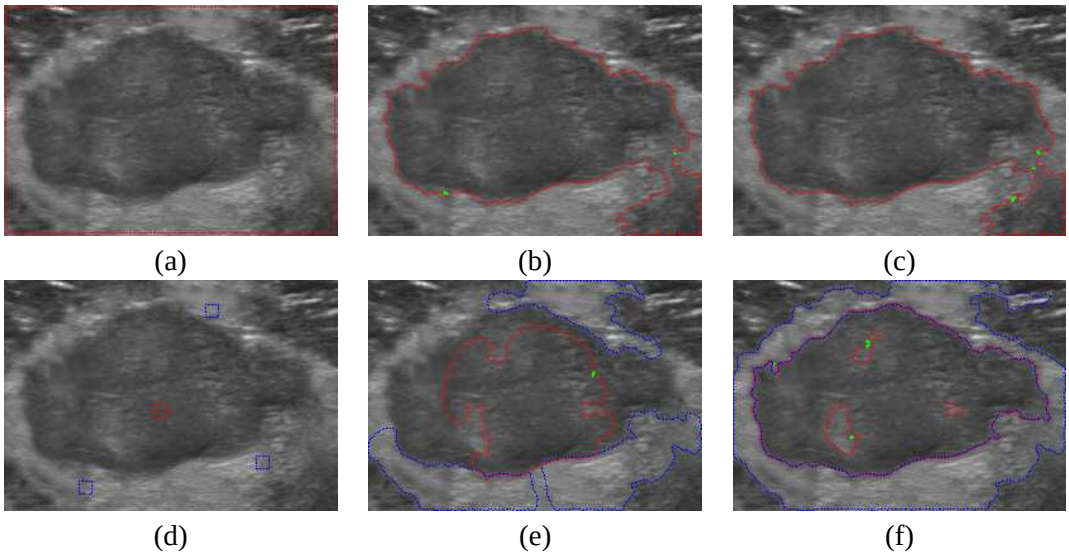


Figure 2. (a–c) Boundary initialization, evolution, and convergence (fail). (d–f) Adaptive initialization, evolution, and convergence (success).

ple growing/shrinking snakes based on the background of the image. Figure 2 shows the segmentation results using the two types of initialization. The boundary initialization could not capture the details of the tumor due to a blurred region of the tumor. With the adaptive initialization, both shrinking and growing snakes correctly captured the tumor.

4 Conclusion

We have shown that our multiple snake algorithm is fast enough for practical applications by measuring the time efficiency of our snake model in successful segmentation experiments of a CAPTCHA challenge data set. With an adaptive initialization technique, our preliminary segmentation results indicates the effectiveness against real medical images with tumors. For the system to be more practical, however, we should automate the initialization procedure in future.

References

[1] M.-F. Auclair-Fortier, D. Ziou, C. Armenakis, and S. Wang. Survey of work on road extraction in aerial and satellite images. Technical Report 247, Departement de mathematiques et d'informatique, Universitede Sherbrooke, 2000.

[2] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1987.

[3] L. D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218, 1991.

[4] L. D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):131–147, 1993.

[5] S.S. Makhanov, R. Marikhu, and M.N. Dailey. Cooperating quadratic snakes for extraction of roads from satellite images. In *Proceedings of MASCOT07, IAC-CNR*, pages 121–130, 2007.

[6] C. Xu and J. L. Prince. Gradient Vector Flow: A new external force for snakes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 66–71, 1997.

[7] M. Rochery, I. H. Jermyn, and J. Zerubia. Higher order active contours. *International Journal of Computer Vision*, 69(1):27–42, 2006.

[8] Drupal CAPTCHA Project. CAPTCHA, 2012. http://drupal.org/project/captcha.

ACIS 2013

Asian Conference on Information Systems 2013

Special Sessions on Onomatopoeia Human Computer Interaction

Product Recommendation Method based on Onomatopoeia Expressing Texture

Ryuichi Doizaki[†] **Saki Iiba**[†] **Takashi Abe**[‡] **Takayuki Okatani**[‡] **Maki Sakamoto**[†]
[†]Graduate School of Informatics and Engineering,
The University of Electro-Communications
[‡]Graduate School of Information Sciences,
Tohoku University
sakamoto@inf.uec.ac.jp

Abstract

With the widespread use of online shopping in recent years, consumer search requests for products have become more diverse. Previous web search methods have used adjectives as input by consumers. However, given that the number of adjectives that can be used to express textures is limited, it is debatable whether adjectives are capable of richly expressing variations of product textures. In Japanese, tactile experiences are easily and frequently expressed by onomatopoeia, such as "*fuwa-fuwa*" which indicates a soft and light sensation. Onomatopoeia are useful for understanding not only material textures but also a user's intuitive, sensitive, and even ambiguous feelings evoked by materials. In this study, we propose a system to recommend products corresponding to product textures associated with Japanese onomatopoeia based on their symbolic sound associations between the onomatopoeia phonemes and the texture sensations. Our system quantitatively estimates the texture sensations of onomatopoeia input by users, and calculates the similarities between the users' impressions of the onomatopoeia and those of product pictures. Our system also suggests products which best match the entered onomatopoeia. An evaluation of our method revealed that the best performance was achieved when the SIFT features, the colors of product pictures, and text describing product pictures were used; Specifically, precision was 66 for the top 15 search results. Our system is expected to contribute to online shopping activity as an intuitive product recommendation system.

Keywords: Product Recommendation Method, Texture, Onomatopoeia, Color, Sound Symbolism

1 Introduction

With the widespread use of online shopping in recent years, consumer search requests for products have become more diverse. Not only are the designs of products, such as their colors and shapes, important factors to consumers, but so are the texture sensations of products, such as tactile sensations. Previous studies have proposed web search methods based on the tactile sensations of the material textures of products [1], [2]. These previous studies propose methods of recommending products to consumers through the input of adjectives such as "soft" or "smooth." However, there is a wide variety of products which can be described by the adjective "soft," such as towels, mouton (sheepskin) fur, rice cakes, and slime, all of which may have soft textures but can differ in their softness. An adjective cannot describe specific tactile information such as the kind or degree of softness of an object. Moreover, the number of adjectives for texture sensations is limited, and thus it is debatable whether such adjectives are capable of richly expressing product texture variations. Therefore, this paper proposes a method of using onomatopoeia based on the symbolic sound associations between onomatopoeia phonemes and texture sensations. Japanese is known to have a large number of onomatopoeic words to express texture sensations. In particular, tactile sensations are easily and frequently expressed by onomatopoeia. The softness of towels, mouton fur, rice cakes, and slime may be expressed, respectively, by "*fuwa-fuwa*" "*moko-moko*" "*mochi-mochi*" and "*funya-funya*". In addition, onomatopoeic expressions such as "*sube-sube*" and "*tsuru-tsuru*" are often used to express the sensation of a smooth surface, and "*neba-neba*" and "*jime-jime*" are often used to refer to sticky sen-

sations. Based on such associations, we propose a new method of estimating texture sensations which can be expressed by onomatopoeia that are suitable as input as consumer search requests.

2 Previous Studies and Purpose of Our Study

2.1 Effectiveness of onomatopoeia in tactile perception

The existence of synesthetic associations between sounds and sensory experiences, referred to as sound symbolism in the literature, has been demonstrated over the decades [3], [4]. It is also known that sensory-sound correspondence can be found not only in words referring to visual shapes, as demonstrated in the landmark studies [4], [5], which identified "mal/mil" and "bouba/kiki" for round and sharp shapes, respectively, but also in words referring to tactile, smell, and taste sensations. [6] has investigated the sound symbolism associations between the phonemes of Japanese sound symbolic words for expressing tactile sensations (i.e., onomatopoeia) and subjective evaluations of comfort or discomfort for touched objects. For example, onomatopoeic words expressing a sense of smoothness often use the consonant /s/ in the first syllable, as in "*sara-sara*" while those expressing roughness often use /z/ in the first syllable, as in "*zara-zara*". [6] also notes that onomatopoeia comprise a class of words whose speech sounds are systematically associated with sensory experiences, and recently they have been used in psychophysical studies as a unique index of sensation categories. [7] has also found that onomatopoeia can be used to differentiate various tactile sensations more easily than adjectives, and that onomatopoeia is thus useful for evaluating minute differences in texture sensations.

2.2 Quantification of affective texture sensations expressed by onomatopoeia

[8] described the sound symbolic relationship between Japanese phonemes and meanings (e.g., the vowel /i/ is associated with sensations of sharpness). [9] quantified the relationship between Japanese phonemes and their perceptions as shown in Table 1, and constructed a model to predict the mental images evoked by two-syllable repetitive onomatopoeia. Based on the model proposed by [9], [10] developed a system to estimate the mental image of various onomatopoeia input by users, as shown in Figure 1.

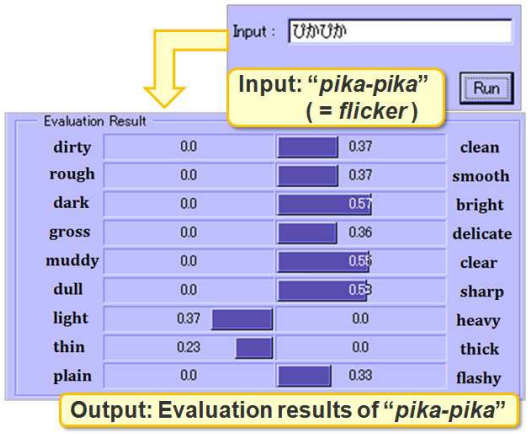


Figure 1. System for evaluating onomatopoeia sensations

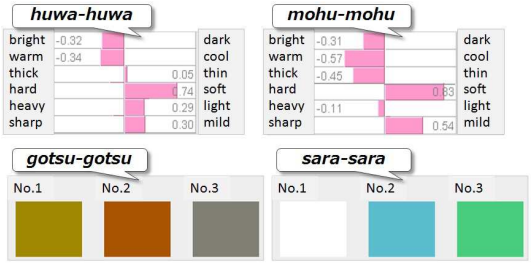


Figure 2. System for evaluating onomatopoeia tactial perceptions

[11] developed a system to evaluate tactile images associated with onomatopoeia based on sound symbolism, as shown in Figure 2. This system quantifies the mental images of input onomatopoeia using 43 adjective scales related to tactile sensations, and then recommends colors which match the impressions associated with the onomatopoeia based on the work of [12], which reported on the influences between the colors of a material and its tactile perceptions.

3 Procedure

3.1 Overview of our system

Figure 3 shows an overview of our system, which quantifies the mental images of textures by analyzing the phonological characteristics comprising the input onomatopoeia, and then recommends products which match the onomatopoeia based on their ranking obtained by

Table 1. Relationship between phonemes and images

Scale	Consonants /t/	/n/	/r/	Sounds voiced sounds	p-sounds
bright / dark	0.09	-0.21	1.21	-1.29	0.69
clear / dull	0.00	-0.45	0.77	-2.47	0.11
heavy / light	0.02	0.42	-0.30	2.58	0.01

calculating the similarities between the texture mental images of the onomatopoeia and the actual texture features of the product pictures. Because our system evaluates mental texture im-

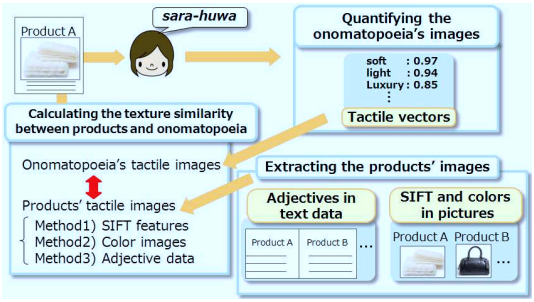


Figure 3. Overview of our system

ages expressed by onomatopoeia based on their phonological characteristics, we began our study by creating a "phoneme image" database. These phoneme image data specify what mental images are associated with each Japanese phoneme in relation to semantic differential (SD) scales. We conducted a psychological experiment to collect the data.

3.2 Stimuli

3.2.1 SD scales

SD scales included 43 pairs of adjectives appropriate for evaluating the design and texture perceptions of objects, such as bright/dark, warm/cool, thick/thin, smooth/rough, and wet/dry. 43 pairs of adjectives were extracted from previous studies related to evaluation of colors or tactile textures. In our method, there are no correlations between 43 adjectives, because our method can calculate the impression values of each adjective independently.

3.2.2 Onomatopoeia

Our method calculates the subjective impressions of onomatopoeia based on the impressions of each phoneme. Therefore, the experimental stimuli are required to include all varieties of Japanese phonemes, that is, basic phonemes (consonants /C/ and vowels /V/) as well as special phonemes (syllabic nasals /N/, choked sounds /Q/, long vowels /R/, and adverbs ending in /ri/) [13]. We created various combinations of sounds to consider the effects of sound order, that is, to investigate whether first-syllable and second-syllable phonemes could evoke different subjective impressions (e.g., between "*kasa*" and "*saka*"). First, we combined all sounds in the Japanese syllabary (from /a/ to /n/), and then created two-syllable expressions (i.e., /aa/, /ai/, . . . , /wan/, /nn/). We obtained a total of 11,075 words, including those made by repeating two-syllable onomatopoeic expressions (e.g., /aa-aa/, /ai-ai/). Moreover, we added 3,509 words with all types of special phonemes, such as /fuwari/ and /peQtari/. Second, we selected 312 words from these 14,584 words which were judged by 3 participants as onomatopoeic expressions for describing texture sensations. We confirmed that the selected 312 words covered every possible kind of phonemeand thus decided to use those 312 words as our experimental stimuli. Table 2 shows some stimuli examples.

3.3 Experiment: Evaluating onomatopoeia images

To create the phoneme image data, we conducted a psychological experiment to investigate what images could represent the 312 onomatopoeic words. Participants were asked to evaluate images for the 312 words (divided into 6 groups of 52 words) against 43 pairs of adjectives across 7-point SD scales, as shown in Figure 4. Onomatopoeic stimuli were displayed on a computer screen randomly for each participant. In total, 78 Japanese undergraduates aged 20 to 24 (51 males and 27 females, spread across 6 groups of 13 people) participated in the experiment.

After obtaining the evaluations (312 words

Table 2. Examples of onomatopoeia stimuli

unyo-unyo	*chiri-chiri*	*mosha-mosha*
gasa-gasa	*tsubu-tsubu*	*yore-yore*
gito-gito	*tsuru-tsuru*	*washa-washa*
gucho-gucho	*teka-teka*	*uNnyuri*
keba-keba	*toro-toro*	*kusyaR*
gotsu-gotsu	*doro-doro*	*zazaQ*
koro-koro	*nayo-nayo*	*zaraR*
gowa-gowa	*nyupu-nyupu*	*subeQ*
zazazaza	*nucha-nucha*	*doroRri*
zashu-zashu	*necha-necha*	*pahuQ*
sara-sara	*fuka-fuka*	*fuRwari*
zara-zara	*bunyu-bunyu*	*moQsa*
jori-jori	*punyu-punyu*	*boko-boko*
sube-sube	*fuwa-fuwa*	*fowa-fowa*
chiku-chiku	*becho-becho*	*muni-muni*
chapu-chapu	*beto-beto*	*washari*

× 43 scales × 13 participants), we calculated the standard deviations of each value and removed 275 evaluations with a standard deviation of 2.0 or over (i.e., 2 %of the total obtained data). By calculating the average values across participants, we thus collected the subjective impressions of all 312 onomatopoeic words (312 words × 43 scales).

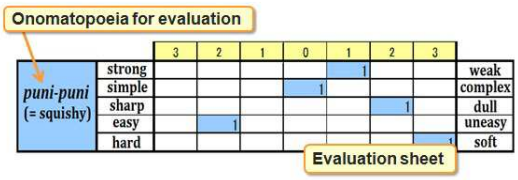


Figure 4. Example of answer sheet in experiment

3.4 Onomatopoeia image estimation model

Our method estimates texture sensations expressed by onomatopoeia based on their phonological characteristics. We designed an "onomatopoeia image estimation model" that quantifies the subjective images of onomatopoeia by calculating a linear sum of the individual phoneme mental images using the following equation (1).

$$\hat{Y} = X_1 + X_2 + \cdots + X_{12} + X_{13} + Const. \quad (1)$$

Here, $\hat{Y}$ represents the mental image estimations of the respective 43 scales for a given onomatopoeia. $X_1 \cdots X_{13}$ show the mental images of the respective phonemes, and $Const.$ represents constant terms. In more detail, $X_1 \cdots X_6$ respectively mean consonant, voiced sound/p-sound, contracted sounds, vowels, semi-vowels and special phonemes in first mora. $X_7 \cdots X_{12}$ respectively mean the same categories in second mora, and X_{13} does repetition.

To obtain the phoneme image data, we applied the onomatopoeia image data used in the experiment to our model, and calculated the images for each phoneme by making use of class I mathematical quantification theory. This theory is a kind of multiple regression analysis using the dummy variable.

For example, the expression "*fuwa-fuwa*" made by repeating /fuwa/, is composed of the first syllable /fu/ (/f/ + /u/) and the second syllable /wa/ (/w/ + /a/). Therefore, the image of "*fuwa-fuwa*" for the "hard/soft" scale is estimated by the following equation. Because the evaluated values are based on 7-points scales, the estimated value 6.28 suggests that "*fuwa-fuwa*" is strongly associated with an image of softness.

$$\begin{aligned} \hat{Y} &= /h/ + /u/ + /w/ + /a/ \\ &\qquad + repetition + Const. \\ &= 0.29 + 0.55 + 0.71 + 0.07 \\ &\qquad + 0.23 + 4.43 \\ &= 6.28 \end{aligned}$$

To evaluate the accuracy of our model, we compared the mental onomatopoeia images estimated by our model with the images of onomatopoeia obtained from participants through our experiment (hereafter, we refer to the former as "estimated images" and the latter as "real images"). We calculated the multiple correlation coefficient $\boldsymbol{R}$ values between the estimated images and real images for the 43 scales. Results showed that for 33 scales, the $\boldsymbol{R}$ values were in the range of 0.8 to 0.9, while for the other 10 scales, the $\boldsymbol{R}$ values were 0.9 or higher. Therefore, we considered our model to be sufficient for estimating onomatopoeia images to be evaluated by humans.

Using equation (1), our system quantifies the onomatopoeia mental images created by consumers for the 43 scales. Thus, we obtain the

vectors for onomatopoeia texture images, that is, the 43 scale values for each image as quantified by the system.

4 Extracting Texture Feelings of Products

In our study, we extracted the design and texture images from actual product pictures and their text descriptions and reviews. To collect pictures to be regarded as products, we used part of the Flickr Material Database (FMD), as described by [14]. The FMD, which was created explicitly for the study of material recognition, consists of 100 pictures of 10 different materials (for a total of 1000 pictures) collected from pictures available on [15]. From these, we chose the five materials of metal, glass, plastic, stone, and wood, for a total of 500 pictures. We also collected 500 text excerpts from the descriptions of the 500 pictures (which we regarded as products in our study). We asked 60 participants to write their impressions of each of the materials in the pictures.

4.1 Method 1: Extracting SIFT features of pictures

The 500 FMD pictures were divided into two sets of 250 pictures each, one set to be used for training and the other for testing. We used the resulting training data for mapping the relationships between the pictures and adjectives. For each set of 250 pictures for training, we manually created sets of related information by the following procedure: (1) two pictures were randomly selected from the set and shown to a participant [16], (2) the participant was asked to answer which picture more represents each adjective. The participant is also allowed to choose "impossible to rank" when the two pictures are considered to be equal in value or when the specified adjective seems unrelated to the pictures.

Using this training data, we calculated the probability of co-occurrence between the adjective scales and the SIFT (Scale-Invariant Feature Transform is an algorithm in computer vision to detect and describe local features in images) feature value for each of the 250 remaining pictures based on the work of [17]. As a result, we obtained the mapping of the relationships between the pictures and adjectives.

Through the above process, we extracted the SIFT feature values for all 500 pictures as texture sensations. These data are regarded as the SIFT vectors for picture texture images, as shown in Figure 5.

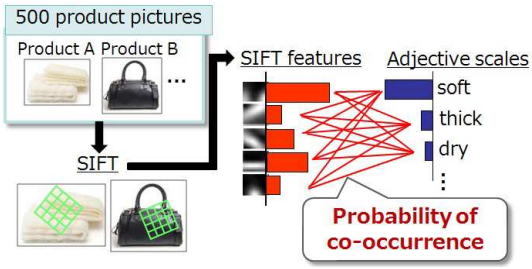


Figure 5. Method 1: Extracting SIFT picture features

4.2 Method 2: Extracting picture color images

In this method, we employed 45 color samples. The color samples were selected by 5 participants as the representative hues and tones among 130 colors, as described in Kobayashi [18].

To collect color image data, we conducted a psychological experiment. Participants were asked to evaluate the images of the 45 colors against 43 pairs of adjectives on 7-points SD scales, as shown in Figure 6. Color stimuli were displayed on a computer screen randomly for each participant. Twenty Japanese undergraduates aged 20 to 24 (10 males and 10 females) participated in the experiment.

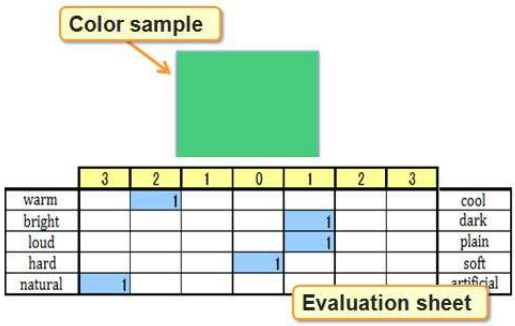


Figure 6. Example of experimental answer sheet

After obtaining the evaluations (i.e., 45 colors $\times$ 43 scales $\times$ 20 participants), we calculated the average values and collected the color image data (45 colors $\times$ 43 scales) as vectors.

We calculated the similarity between the obtained vectors for onomatopoeia texture images

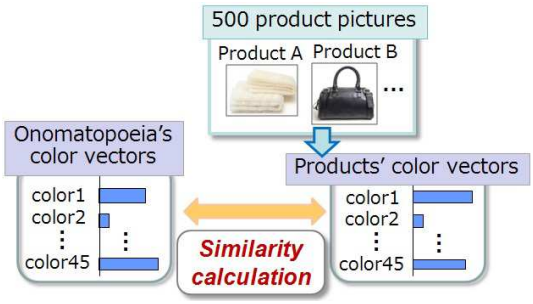


Figure 7. Method 2: Extracting color vectors of pictures

and color image vectors for each color. Furthermore, for the 500 pictures, we calculated the color frequency in each pixel, and each picture was assigned 45 color feature values. Using the process above, we obtained the color feature values as the picture color vectors for the pictures, as shown in Figure 7

4.3 Method 3: Extracting texture sensations from text data

As noted previously, we collected 500 text excerpts describing the products and then asked 60 subjects to write down their subjective impressions of the materials shown in the pictures freely. We assumed the text data to represent product reviews, and then used the collected data in the following analysis.

We extracted the adjectives included in the 43 adjective scales from the collected text data. We used the concept dictionary introduced in [19] and an online thesaurus [20] to extract the relevant adjectives.

We calculated the frequency of the adjectives appearing in the 500 text excerpts. We calculated the importance of the adjective scale results for each product picture using TF-IDF (Term Frequency Inverse Document Frequency is a numerical statistic which reflects how important a word is to a document in a collection or corpus). We obtained the importance of the adjectives in relation to the pictures as adjective vectors for picture texture perceptions.

4.4 Calculating the texture similarity between products and onomatopoeia

For each method mentioned above, product pictures were ranked according to their degree of matching with respect to the onomatopoeia by calculating the cosine similarity between the vectors of onomatopoeia texture perceptions and the vectors of picture texture perceptions.

The vectors of onomatopoeia texture perceptions, calculated by equation (1), comprised 43 values evaluated based on each adjective scale. Here, if a texture image vector of an onomatopoeia f is $\boldsymbol{k}(f)$, and those of a product picture t, which is recommended by each method, is $\boldsymbol{k}(t)$, then the cosine similarity between the two vectors ($\cos(\boldsymbol{k}(f), \boldsymbol{k}(t))$) is given by following equation (2). For all product pictures, we calculated the onomatopoeia cosine values, and recommended the product picture having higher cosine values than the other pictures.

$$\cos(\boldsymbol{k}(f), \boldsymbol{k}(t)) = \frac{\boldsymbol{k}(f) \cdot \boldsymbol{k}(t)}{|\boldsymbol{k}(f)||\boldsymbol{k}(t)|} \qquad (2)$$

We implemented a product recommendation system based on each method as a Java program. Figure 8 shows examples of recommendation results based on each of the three methods.

Figure 8. Recommendation results using each method

5 Evaluation

To verify the efficiency of our system, we used 60 pictures of the 500 product pictures for evaluation, with each picture having one text excerpt obtained in Method 3. Ten subjects were asked to provide responses for onomatopoeia expressing texture perceptions of the materials in each picture. For this task, (1) we chose onomatopoeia corresponding to 10 or more pictures, and also (2) selected onomatopoeia answers provided by at least 3 subjects. Those that satisfied both of these conditions included the following six onomatopoeia: *tsuru-tsuru*, *zara-zara*, *teka-teka*, *sube-sube*, *pika-pika*, *gotsu-gotsu*. These

Table 3. Precision $\boldsymbol{P}$ and reproducibility $\boldsymbol{R}$ in top $\boldsymbol{N}$ search results

Top N search results	**Method 1 SIFT features**		**Method 2 Color images**		**Method 3 Adjectives**		**All Methods 1 + 2 + 3**	
	$\boldsymbol{P}$	$\boldsymbol{R}$	$\boldsymbol{P}$	$\boldsymbol{R}$	$\boldsymbol{P}$	$\boldsymbol{R}$	$\boldsymbol{P}$	$\boldsymbol{R}$
5	0.73^{*1}	0.18^{*2}	0.37	0.08	0.63	0.17	0.63	0.15
10	0.63	0.3	0.37	0.17	0.63	0.34^{*2}	0.67^{*1}	0.33
15	0.64	0.48	0.38	0.26	0.58	0.46	0.66^{*1}	0.5^{*2}
20	0.6	0.6	0.36	0.34	0.58	0.6	0.63^{*1}	0.62^{*2}
25	0.54	0.66	0.36	0.42	0.56	0.72^{*2}	0.57^{*1}	0.72^{*2}
30	0.51	0.75	0.35	0.5	0.52	0.8	0.54^{*1}	0.82^{*2}

were thus categorized as the most accurate onomatopoeia to describe the pictures.

Using the degree of precision (percentage of correct data in the top $\boldsymbol{N}$ search results) and reproducibility (percentage of correct data in the top $\boldsymbol{N}$ search results in among correct data), we compared the performance of each method. In cases of the integration of multiple methods, we calculated the average similarities across all methods. We calculated the precision and reproducibility values for each of the six onomatopoeia, and then obtained the average precision $\boldsymbol{P}$ and average reproducibility $\boldsymbol{R}$ for each, as shown in Table 3. In the table, "*1" means the highest precision value, and "*2" does the highest reproducibility value for the top $\boldsymbol{N}$ search results.

As Table 3 shows, when we integrated all three methods, we obtained the highest values for both precision and reproducibility. We thus found that our system shows the best performance using SIFT features (for product pictures), colors (for product pictures) and adjectives (for text data). When integrating all three methods, precision reached 66 %, and reproducibility reached 82 % in more than top 15 search results. This result shows that our system succeeded in recommending products which cognitively fit consumers' texture perceptions expressed by onomatopoeia.

6 Conclusion

In this study we proposed a system for recommending products which match texture perceptions expressed by onomatopoeia based on sound symbolism. Our system quantitatively estimates the texture perceptions of input onomatopoeic expressions by users, and then calculates the similarities between the texture perceptions expressed by those onomatopoeia and those of product pictures. Finally, our system suggests products to the user which match the perception or sensation expressed by the onomatopoeia.

The evaluation of our system shows that when integrating all three of the proposed methods, the system shows the best performance. Moreover, we were able to verify the validity of our system.

In our system, we chose the five materials of metal, glass, plastic, stone, and wood as search categories. By expanding these product material categories in future work, we expect that our system will correspond to more and more materials required by consumers as they perform online search activities.

To these ends, in future work we aim to examine the practical uses of our system in actual web search situations.

Acknowledgment

This study was supported by Grant-in-Aid for Scientific Research on Innovative Areas "Shitsukan" (No. 23135007) from MEXT, Japan.

References

[1] E. Yano, Y. Kitano, E. Sueyoshi, I. Shinohara, P. Sineenard, and T. Kato. Development of recommendation system with anonymous kansei model. *Information Processing Society of Japan: Database*, 44(SIG8(TOD18)):46–54, 2003.

[2] Y. Murakami, S. Nakamura, and S.Hashimoto. An article search system that learns the kansei of the user. *Proceedings of the 72th National Convention of the Information Processing Society of Japan*, 2:451–452, 2010.

[3] W. Köhler. *Gestalt Psychology*. Liveright Pub., New York, 1929.

[4] E. Sapir. A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3):225–239, 1929.

[5] V. S. Ramachandran and E. M. Hubbard. Synaesthesia a window into perception. *Thought and Language Journal of Consciousness Studies*, 8(12):3–34, 2001.

[6] J. Watanabe, T. Hayakawa, S. Matsui, A. Kano, Y. Shimizu, and M. Sakamoto. Visualization of tactile material relationships using sound symbolic words. *P. Isokoski and J. Springare (Eds.): EuroHaptics 2012*, LNCS 7283(Part II):175–180, 2012.

[7] M. Sakamoto and J. Watanabe. Effectiveness of onomatopoeia representing quality of tactile texture - a comparative study with adjectives. *Proceedings of the The 13th Annual Conference of the Japanese Cognitive Linguistics Association*, pages 242–245, 2012.

[8] S. Hamano. *The Sound-symbolic System of Japanese*. University of Florida, Gainesville, 1986.

[9] N. Fujisawa, F. Obata, M. Takada, and S. Iwamiya. Impression of auditory imagery associated with japanese 2-mora onomatopoeic representation. *Journal of the Acoustical Society of Japan*, 62(11):774–783, 2006.

[10] Y. Shimizu and M. Sakamoto. Creation support system for japanese onomatopoeia based on sound symbolism. *Proceedings of the The 26th Annual Conference of the Japanese Society for Artificial Intelligence*, 2N1-OS-8c-4:1–3, 2012.

[11] S. Iiba, Y. Shimizu, and M. Sakamoto. A method to select colors appropriate for tactile onomatopoeia based on sound symbolism. *Proceedings of the International Workshop on Modern Science and Technology (IWMST2012)*, pages 127–131, 2012.

[12] L. T. Maloney and H. Brainard. Perception of color and material in complex scenes. *Journal of Vision*, 10(9):1, 2010.

[13] H. Kakehi, I. Tamori, and L. C. Schourup. *Dictionary of Iconic Expressions in Japanese*. Mouton de Gruyter, Berlin, New York, 1996.

[14] C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz. Exploring features in a bayesian framework for material recognition. *In Proc. CVPR*, 2010.

[15] Flickr material database. `http://people.csail.mit.edu/celiu/CVPR2010/FMD/`.

[16] Amazon mechanical turk. `https://www.mturk.com/mturk/welcome`.

[17] T. Abe, T. Okatani, and K. Deguchi. Recognizing surface qualities from natural images based on learning to rank. *21st International Conference on Pattern Recognition (ICPR 2012)*, pages 3712–3715, 2012.

[18] S. Kobayashi. *Color Image Scale*. Kodansya, Tokyo, 2001.

[19] T. Yamaguchi. *Japanese Thesaurus - Categorized word dictionary-*. Taishukan, Tokyo, 2003.

[20] Wordnet. `http://nlpwww.nict.go.jp/wn-ja/index.ja.html`.

Method of evaluating metal textures by the sound symbolism of onomatopoeia

Junya Yoshino[†] **Akira Yakata**[‡] **Yuichiro Shimizu**[†] **Masaharu Haginoya**[*] **Maki Sakamoto**[†]
[†]Graduate School of Informatics and Engineering,
The University of Electro-Communications
[‡]Faculty of Electro-Communications,
The University of Electro-Communications
[*]Calsonic Kansei Corporation
sakamoto@inf.uec.ac.jp

Abstract

In recent years many products, including many electric appliances, are starting to be made of imitation metals, which are lighter in weight and lower in cost than real metals. However, imitation metals can sometimes look and feel cheaper. This study proposes that differences in mimetic and onomatopoeia associated with imitation and real metals can be used in metal texture design to make imitation materials look and feel more like real materials. Japanese is a language known to have a large number of mimetic and onomatopoeia for expressing textures. For example, such words expressing a sense of smoothness often use the consonant /s/ in their first syllable, as in "*sara-sara*". We conducted psychological experiments where participants were asked to look at a pair of imitation and real materials without touching them and respond with onomatopoeia that they felt were associated with them. Participants responded significantly more onomatopetic words for real materials than for imitation materials. Furthermore, we analyzed ono-matopoeia obtained from both real metals and imitation metals using a system that quantifies the phonological information of onomatopoeia. A comparison of the onomatopoeia evoked by real metals and those evoked by imitation metals revealed differences in the affective features of the two types of metals.

Keywords: Metal Texture Design, Real Metal, Imitation, Onomatopoeia

1 Introduction

In recent years, many products, including electric appliances and car interiors, are increasingly made of imitation metals, which are lighter in weight and lower in cost than real metals. While at first glance imitation metals may appear similar to real metals, they can sometimes look and feel cheaper. In product development, the field of metal texture design aims to make imitation metals look as similar to real metals as possible. However, the reason why imitation materials often feel different from real metals even if given the same texture design remains a mystery, and it is still unclear what kinds of design are effective to make such textures more realistic. Humans can infer the properties of materials through the five senses; namely, vision, touch, smell, taste, and hearing. The present paper focuses on vision. Many studies have investigated how humans interpret visual input to, for example, estimate the properties of surfaces, how to differentiate peaches and nectarines, or how to tell the difference between unfinished and polished wood [1] , [2]. There are many visual attributes that help us to distinguish different surface materials, such as lightness, color, and texture. For example, [1] pointed out that one of the basic characteristics of image statistics, surface perception —the distribution of luminance values in an image, or the "skew"—is highly correlated with subjective judgments of gloss and lightness.

However, while such findings reported by previous studies have focused mainly on material perceptions through experiments using pictures, there is still little research on the emotional or sensuous perceptions of texture. [3] was one such study that focused on the emotional perception of metal textures in an attempt to analyze the features of human texture perception. However, the stimuli used in their experiment were, nonetheless, images of textures.

To pursue a new method of capturing the emotional and sensuous aspects of material per-

ception, this study focuses on onomatopoeia as sound expressions that can symbolize various textures. The existence of synesthetic associations between sounds and sensory experiences, known as sound symbolism, has been demonstrated [4], [5]. It is also known that a sensory-sound correspondence can be found not only in words referring to visual shapes, such as in the examples "mal/mil" and "buba/kiki" for describing round and sharp shapes, respectively, as noted in the landmark studies of [5] and [6], but also in words referring to tactile, smell, and taste sensations [7], [8] . Japanese is known to have a large number of onomatopoeia; for example, [9] describes the sound symbolic associations between the phonemes of Japanese onomatopoeia and their particular meanings, as shown in Table 1.

[8], [10] investigated the sound symbolic associations between the phonemes of Japanese onomatopoeia for expressing tactile sensations and the subjective evaluations of comfort/discomfort for touched objects. For example, onomatopoeia ex-pressing a sense of smoothness often use the consonant /s/ in their first syllable, as in "*sara-sara*" while those expressing roughness often use /z/ in their first syllable, as in "*zara-zara*" However, no study has attempted to use onomatopoeia to explore visual differences between real and imitative materials with similar texture designs. The purpose of this study is therefore to investigate the effectiveness of onomatopoeia in the visual measurement of metal textures and explore the possibility that differences in the onomatopoeia associated with imitation and real metals can be used for metal texture design with the aim of making imitation materials look and feel more like real materials.

2 Experiment

2.1 Materials and design

In this experiment, we aimed to compare the numbers and types of onomatopoeia associated with real metals with those associated with imitation metals. [8], [10] notes that onomatopoeia is effective in understanding the sensibilities of human texture perception, such as likes or dislikes and comfort or discomfort; thus we hypothesize that real metals are more easily associated with onomatopoeia than metal-like materials, namely imitation metals.

A previous study by [11] on the texture of lacquer found that expert participants evaluated the characteristics of materials more accurately than non-experts. Therefore, we first chose 20 experts to be participants in our experiment; specifically, they were 18 men and two women who had worked for over 5 years in professions dealing with metal surface design. Then we chose 30 non-experts (18 men and 12 women) to comprise a comparison group. The experiment was per-formed using a box set up with a light color temperature of 5500 K, and in which a pair of real and imitation metals was placed, as shown in Figure 1.

Figure 1. Experimental setup

The experimental stimuli were 15 pairs of real metals and imitation metal-like materials with the same surface texture designs shown in Figure 2. In Figure 2, the material shown in the right side of each image is the imitation, and the material on the left side is the real metal. However, the placement of the materials was randomized for each participant. The experimental stimuli were made by Calsonic Kansei Corporation, Saitama prefecture, Japan.

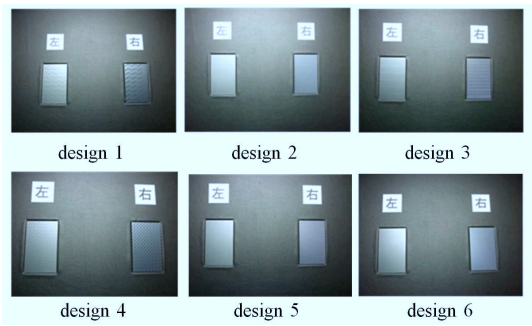


Figure 2. Examples of experimental stimuli

Table 1. Examples of sound symbolic associations between Japanese phonemes and meanings

Vowel		Consonant	
/i/	line, extended in a straight line	**/h/**	softness, weakness
/a/	flat, spreading, large surface	**/s/**	smoothness, serenity, tranquility

2.2 Procedure

Participants were given instructions on the procedure and a brief description of onomatopoeia, along with some examples of Japanese onomatopoeic expressions that are typically used to describe things other than metal (so there would be no influence on the actual experiment). Participants were asked to sit in front of the box with their forehead placed on the board attached to the top of the box, as shown in Figure 3. Then, they were asked to observe a pair of objects consisting of one real metal and one imitation metal-like material, and then provide as many onomatopoeic expressions they could associate with the objects as they could during a 20-second period of response time. The participants were instructed not to touch the materials, because touching can reveal non-visual differences between real and imitation metals.

Figure 3. Image of experimental trial

2.3 Results

The average number of onomatopoeia obtained in response to the real metal stimuli was 1.50 (1.36 from the experts and 1.59 from the non-experts), while the average number of onomatopoeia obtained in response to the imitation metal stimuli was 1.07 (0.94 from the experts and 1.16 from the non-experts). We conducted a t-test (two-tailed, 5% confidence level) to examine whether there was a significant difference between the average numbers of onomatopoeia for the real metals and the imitations.

A two-tailed t-test (both sides, 5% confidence level) revealed a significant difference between the two averages: t(49)=8.87, $p < .01$ (total), t(19)=6.61, $p < .01$ (experts) and t(29)=6.28, $p < .01$ (non-experts). This result suggests that real metals were more easily associated with onomatopoeia than imitation metal-like materials. This result further suggests that even if two objects have the same surface design, real metals are more easily associated with emotional texture perceptions and thereby more easily evoke a variety of onomatopoeia, which are presumed to be directly associated with both intuitive and emotional perception.

Furthermore, we compared the types of onomatopoeia obtained from the non-experts with those obtained from the experts. We regarded the word given most frequently by each group of participants to describe each material as the representative word for the material for that group. Table 2 shows the representative onomatopoeia obtained for each material. Tactile onomatopoeia such as "*zara-zara*" (describing a rough texture) were obtained from non-experts more than from experts, while experts responded with more visual-oriented onomatopoeia such as "*teka-teka*", which expresses brightness. Therefore, we conducted a statistical analysis to determine whether tactile onomatopoeia were more frequently provided by non-experts than experts. Results of the analysis showed that tactile onomatopoeia were provided by non-experts significantly more frequently than by experts: χ 2(1,N=75)=16.835, $p < .01$. This result suggests that while metal texture design experts tend to focus on the visual design or qualities of metal surfaces, non-experts —that is, lay persons —tend to perceive material properties through experiences closer to the sense of touch rather than vision.

Table 2. Representative onomatopoeia associated with each material for each participant type

	Non-experts		Experts	
Design	**real metal**	**imitation**	**real metal**	**imitation**
1	zara-zara	tsuru-tsuru	une-une gira-gira teka-teka	sara-sara
2	boko-boko	tsuru-tsuru	kira-kira gira-gira	sara-sara

3 Analysis by New Evaluation Method

3.1 Method overview

We analyzed the information expressed by the onomatopoeia given in Table 2 using a method that estimates the texture feelings expressed by onomatopoeia based on their phonological characteristics. We designed an onomatopoeic image estimation model that quantifies images evoked by onomatopoeia by calculating a linear sum of the phoneme evocations (i.e., perceptions) using the following equation (1).

$$\hat{Y} = X_1 + X_2 + \cdots + X_{12} + X_{13} + Const. \quad (1)$$

Here, $\hat{Y}$ represents the estimation of perceptions of 43 measurement scales for a given onomatopoeia. X1-X13 represent the perceptions of the respective phonemes, and Const. represents constant terms. To obtain the phoneme perception data, we applied the onomatopoeia image data from our experiment to our model, and calculated the perceptions of each phoneme using a class I mathematical quantification theory. This theory is a kind of multiple regression analysis using the dummy variable.

We conducted a psychological experiment to create the phoneme perception data. Participants were asked to evaluate their subjective impressions of 312 onomatopoeia (six groups of 52 words) against 43 pairs of adjectives in 7-point semantic differential (SD) scales. Onomatopoeic stimuli were randomly displayed on a computer screen for each participant as shown in Figure 4. In total, 78 Japanese undergraduates aged 20 to 24 (51 males and 27 females, divided into six groups of 13) participated in the experiment.

Results of the experiment included 174,408 evaluations (312 words × 43 scales × 13 participants). We then calculated the average evaluated values for the onomatopoeic perception data

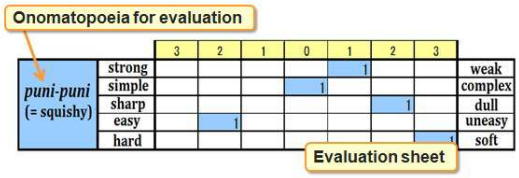


Figure 4. Image of experimental trial

across the participants (312 words × 43 scales). We then applied the onomatopoeic perception data to our model and calculated the perception values of each phoneme by making use of a class I mathematical quantification theory.

For example, the expression "*fuwa-fuwa*" made by repeating /fuwa/, is composed of the first syllable /fu/ (/f/ + /u/) and the second syllable /wa/ (/w/ + /a/). Therefore, the image of "*fuwa-fuwa*" for the "hard/soft" scale is estimated by the following equation. Because the evaluated values are based on 7-points scales, the estimated value 6.28 suggests that "*fuwa-fuwa*" is strongly associated with an image of softness.

$$\begin{aligned} \hat{Y} &= /h/ + /u/ + /w/ + /a/ \\ &\qquad + repetition + Const. \\ &= 0.29 + 0.55 + 0.71 + 0.07 \\ &\qquad + 0.23 + 4.43 \\ &= 6.28 \end{aligned}$$

To evaluate the accuracy of our model, we compared the perception values estimated by our model with the real values for the 312 onomatopoeic stimuli obtained from our participants in experiment. (Hereafter, we refer to the former as "estimated values" and the latter as "real values".) We calculated the multiple correlation coefficients between the estimated values and the real values for the 43 scales. Results showed that

for 33 scales, the multiple correlation coefficient values were in the range of 0.8 to 0.9, while for the other 10 scales the values were 0.9 or higher. We thus considered our model sufficiently valid for estimating human onomatopoeic perceptions.

3.2 Analysis results

Using our method, we calculated the information expressed by onomatopoeia words such as those shown in Table 2, all of which were obtained in the psychological experiments described in Section 2 for the 43 scales. Table 3 shows examples of the output by our method.

We analyzed the difference between the evaluation values of onomatopoeia words obtained from real metals and imitations of each design included in the experiment. We calculated the absolute value of each scale and found that texture perceptions between the real metals and their imitations were the most different for the following evaluation scales: [smooth/rough], [heavy/light], [clean/dirty], and [wet/dry].

3.3 Second Evaluative Experiment

Because our findings revealed that one of the most different scales between the real and imitation metal perceptions was [smooth/rough], we conducted another experiment focusing on Design 6, which had been evaluated differently on the scale of [smooth/rough]. Our method had estimated that humans would evaluate the real metal as being rougher than the imitation, even when the texture design was the same. Namely, Design 6 features geometrical pattern of 0.2 mm in depth for both the real and imitation metal stimuli. Therefore, to make the imitation metal stimulus look more similar to the real metal stimulus, we increased the roughness of the imitation metal by increasing the depths of the geometrical pattern to 0.6 mm and 0.4 mm. (The design itself was created by Calsonic Kansei Corporation.)

We conducted an experiment using the real Design 6 metal and the newly created, rougher imitation metal. The procedure of the experiment was the same as that described in Section 2.2. Participants were asked to sit in front of the box with their forehead placed on the board attached to the top of the box, and asked to compare among a sample of the real metal (placed in the middle), the original imitation sample (placed to either the right or left side), and the newly created imitation sample (placed on the opposite side not occupied by the original imitation), as shown in Figure 5. Participants were asked to answer which object on the sides seemed closer to the real metal object in the middle. We conducted this procedure for imitations with depths of 0.2 mm, 0.4 mm, and 0.6 mm.

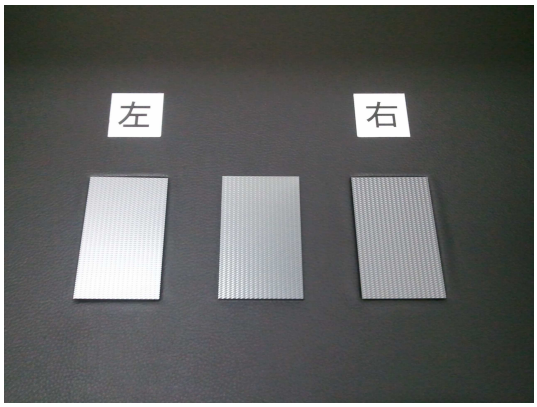


Figure 5. Image of experimental trial

3.4 Evaluation results

Results showed that all participants responded that the new metal-like material with a depth of 0.6 mm seemed most similar to the real metal. This result indicates that by comparing results of the experimental outputs between onomatopoeia associated with real metals and those with metal-like materials, we can better understand how to create imitation metals to look more like real metals by modifying their texture design as indicated by the experimental analysis.

4 Conclusion

This study suggests that differences in onomatopoeia associated with real and imitation metals can be used in metal texture design to create imitation materials that look and feel more like real materials. Through psychological experiments in which participants were asked to visually evaluate a pair of imitation and real materials without touching them and respond with onomatopoeia associated with such materials, we showed that real materials were significantly more easily associated with more onomatopoeia than were imitation materials. This result indicates that even if imitation metals have similar surface designs as real metals, real metals are still more easily associated with emotional

Table 3. Examples of output from our system

Non-experts（style 1）	**Real metal(R)**	**Imitation(I)**	**R − I**
scales（1 ←:→ 7）	**zara-zara**	**tsuru-tsuru**	**difference**
warm : cold	4.76	3.88	0.88
wet: dry	5.15	3.47	1.68
smooth : rough	4.90	2.75	2.15

texture perceptions and as such more easily evoke onomatopoeia, which have been found to be directly associated with intuitive and emotional perception. This finding suggests that such differences in onomatopoeia associated with imitation and real metals can be used to inform metal texture design to make imitation materials look and feel more like real materials, while also providing useful insights to metal texture design experts.

Furthermore, we analyzed onomatopoeia obtained from responses to observing real and imitation metals using a method that quantifies onomatopoeic phonetic (phoneme) information. A comparison of outputs between the onomatopoeia evoked in response to real metals and those evoked in response to imitation metals revealed a difference in the affective features of the two types of metals. Based on these differences, we created new metal-like materials. After testing these new materials, we confirmed that we succeeded in making metal-like materials that were perceived as being more similar to their real metal counterparts.

In our future work, we plan to apply this method to various other materials, such as pearl, glass, and stone.

Acknowledgment

This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas "Shitsukan" (No. 23135510) from MEXT, Japan

References

[1] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson. Image statistics and the perception of surface qualities. *Nature*, 447:206–209, 2007.

[2] D. E. Ho, K. Imai, K. Gray, and E. A. Stuart. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15:199–236, 2007.

[3] T. Furukawa, N. Yata, and Y. Manabe. Subjective evaluation of metal like ap-pearance for estimating the determinants of material perception. *Journal of Meeting of Special Interest Group on "Foundations of Visual Information"*, 11:38–41, 2012.

[4] W. Köhler. *Gestalt Psychology*. Liveright Pub., New York, 1929.

[5] E. Sapir. A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3):225–239, 1929.

[6] V. S. Ramachandran and E. M. Hubbard. Synaesthesia a window into perception. *Thought and Language Journal of Consciousness Studies*, 8(12):3–34, 2001.

[7] J. Watanabe and M. Sakamoto. Sound symbolic relationship between onomatopeias and emotional evaluations in taste and touch. *9th International Symposium on Iconicity in Language and Literature*, 2013.

[8] J. Watanabe, T. Hayakawa, S. Matsui, A. Kano, Y. Shimizu, and M. Sakamoto. Visualization of tactile material relationships using sound symbolic words. *P. Isokoski and J. Springare (Eds.): EuroHaptics 2012*, LNCS 7283(Part II):175–180, 2012.

[9] S. Hamano. *The Sound-symbolic System of Japanese*. University of Florida, Gainesville, 1986.

[10] J. Watanabe, A. Kano, Y. Shimizu, and M. Sakamoto. Relationship between judgments of comfort and phonemes of onomatopoeia in touch. *The Virtual Reality Society of Japan*, 16(3):367–370, 2011.

[11] W. Lee, M. Sato, T. Asami, H. Oyabu, and S. Tominaga. Relation of light reflection to visual texture on black coating films of

urushi (japanese lacquer) and various synthetic resins. *The Color Science Association of Japan*, 26(4):236–247, 2002.

A Basic Study of Hot-and-cold Blower for Realizing Onomatopoeias

Yasushi Hamamura† **Mitsunori Matsushita**‡
†Graduate School of Informatics, Kansai University
‡Faculty of Informatics, Kansai University
k220843@kansai-u.ac.jp

Abstract

The goal of this study is to develop a system for virtually experiencing onomatopoeias, which are echoic or mimetic words. As compared to other languages, Japanese features a large number of onomatopoeias, which serve as a hindrance to foreigners learning Japanese. In order to overcome this hindrance, we present tactile information corresponding to onomatopoeia and prompt its physical understanding. We focus on temperature and wind senses and propose a hot-and-cold blower, which consists of a temperature generator and a blower. Further, we conduct three experiments to verify the effectiveness and correlativity of the proposed system.

Keywords: onomatopoeia, temperature sense, wind sense, tactile interface

1 Introduction

Over the last few decades, the number of people learning Japanese overseas has increased dramatically, going from a mere 130,000 in 1979 to over 3,650,000 in 2009, according to the Japan Foundation[1]. One of the big challenges for these new learners of Japanese is understanding the language's peculiar use of onomatopoeia (i.e. echoic or mimetic words) that sensually express information like sound, texture or movement. The meanings of onomatopoeic words for brightness (pikapika, donyori, etc.) or hardness (fuwafuwa, kachikachi, etc.), for example, are highly abstract and hard to quantify, but because such words are common in Japanese, gaining a good understanding of them is essential to grasping the broader semantics of the language.

A number of efforts have focused on teaching Japanese onomatopoeia more effectively. An online learning support system called ONOMATOPENARAI[2], for example, helps individual learners by illustrating how to use selected onomatopoeic words through example sentences, similar expressions, etc. Another example is the Onomatopoeia Sound Database[3], which provides learners with sound data corresponding to various onomatopoeic vocabularies.

Aside from these self-learning systems, several tools and applications that integrate onomatopoeic keywords have emerged in recent years. Examples include Onomatoperori[4], a cooking application that offers recipe recommendations based on onomatopoeic input, and Onomatopen[5], a painting application that changes brush styles according to spoken onomatopoeic words. The sensual contexts of these applications can make the meaning and application of onomatopoeic keywords more accessible to users.

One range of Japanese onomatopoeia that has not been addressed by such applications is tactile onomatopoeia. This is hardly surprising, since existing applications depend heavily on visual and auditory input and output. Nonetheless, if a system could integrate tactile information to convey the meaning of tactile onomatopoeia directly, it could provide a major boost to intuitive learning of Japanese.

In this study, we propose a system that produces tactile information based on onomatopoeic keywords. Focusing on the user's sense of temperature and wind quality, we integrate a "hot-and-cold blower" device to produce experiences associated with those keywords.

2 Related Work

Systems integrating tactile information have generally used it in one of two ways: (1) as input/output device for system, or (2) as a reinforcement of visual and auditory information. In this section, we limit our discussion to systems that use temperature and wind interfaces.

2.1 Temperature Interface

In this section we take "Themotaxis"[6] and "Thermoesthesia"[7] as examples of the systems using temperature information. Thermotaxis is

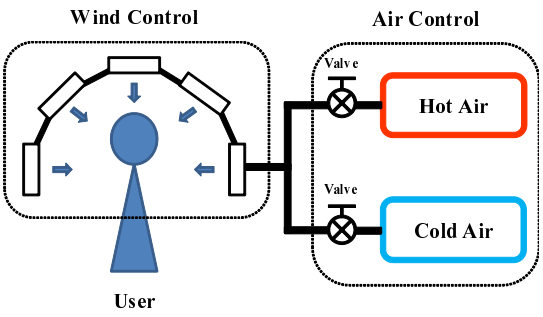


Figure 1. Initial Prototype Design

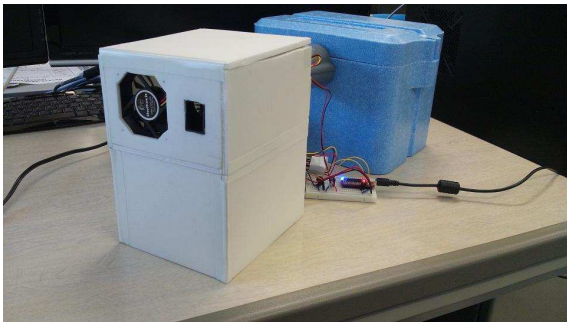


Figure 2. Experimental Device Setup

a system for conducting variable temperature information to users through an earmuff-type wearable device. This system uses temperature information as an input. To control human behavior and their position implicitly by presenting environmental information, two types of Thermotaxis were invented. One, "Thermotaxis: defined," presents thermal sensation and thus influences standing position of users. The other, "Thermotaxis: society," creates thermal sensory spots around people and thus encourage uses to gather together or come in comfortable personal distance.

Another system, Thermoesthesia, uses temperature information as a reinforcement of visual and auditory information. The system is designed to let users feel the temperature of warm or cool displayed objects, by directly touching the display surface.

Such interfaces are understood intuitively by most users, even when they convey relatively abstract or emotional meanings (e.g. joy or sorrow) using temperature (e.g. warmth and chill). When visual components are added (e.g. fire and ice) the intuitive is magnified even further.

2.2 Wind Interface

Most systems that make use of wind information do so in combination with visual information, and many integrate audio as well.

The Wind-Surround System[8] uses a set of blowers arranged in circular fashion around the user to reproduce wind environments, while a head-mount display and earphones reproduce corresponding visual and an auditory output. The system is intended to give a more realistic expression of virtual reality (e.g. of a storm, winged flight, etc.)

3 Proposed System

3.1 Basic Design

The hot and cold air currents that a human being feels in everyday life can be described using several key parameters: strength, direction, duration, temperature, and humidity. Our hot-and-cold blower system was designed to reproduce as many of these parameters as possible, so as to cover the greatest number of related onomatopoeic expressions.

We can divide the above parameters into two classes: (1) wind parameters (strength, direction and duration) and (2) air parameters (temperature and humidity). These classes correspond to the subsystems for wind-control and air-control illustrated in Figure 1. In the wind-control subsystem, wind strength, direction, and duration are controlled by adjusting the output of blower fans installed at different positions around the user. In the air-control subsystem, temperature and humidity are controlled by adjusting the apertures of blower-integrated air tanks with different levels of heat and moisture.

In this paper, we focused on two parameters, the strength of wind and temperature, which are supposed to be especially important parameters in hot and cold winds. For this purpose, we made the device which connects wind-control and air-control subsystems through a single short pipe.

3.2 Temperature Generator

Our hot-and-cold blower system, pictured in Figure 2, consists of a temperature generator that feeds temperature-adjusted air to a wind generator, which then blows the air at an adjusted strength.

Many systems that produce temperature-adjusted output make use of Peltier devices, as

they are small, produce little noise or vibration, and can control temperature with great accuracy. Unfortunately, Peltier devices would not produce enough heat to sufficiently alter the temperature of air flow used in our system. For this reason, we chose instead to control the temperature of an insulated chamber made of expanded polystyrene casing, using coolants and heating elements. Specifically, we introduced a 1 kilogram mass of dry ice as our coolant and a ceramic fan heater as our heating element.

3.3 Wind Production Device

Our wind-production device used two DC fans: an Owltech OWL-FY0615M as a propeller, and a Sunon GB1205PKV1-8AY.GN as a blower. The propeller fan produces greater air flow than the blower, but a lower static pressure. Using the two in combination, we can reproduce a wider range of wind dynamics.

The output of the two fans is under PWM (Pulse Width Modulation) control via a microcontroller (Gainer Mini) connected by USB. In this way, we can command 256 different levels of wind output through the PC software platform known as "Processing." Since the Gainer Mini microcontroller provides only 5V of electrical power and the fans both require 12V, we incorporated a TOSHIBA transistor (model 2SK2231) to provide additional voltage. The fans will still cease to operate when the PWM value dips below a certain threshold, but they can still reproduce a large range of wind strengths.

3.4 Onomatopoeia for Presentation to Test Subjects

To test our system, we chose a total of 17 wind-related (e.g., pyu-, buo-, etc.) and 13 temperature-related (e.g., hinyari, pokapoka, etc.) words from the Onomatopoeia Dictionary[9]. We then analyzed these words with respect to some parameters (wind, temperature, humidity, contact area, etc.). Since the current prototype of our system only reproduces wind strength and temperature, we left other parameters to the imagination of the subjects and didn't take into consideration.

4 Experimental Evaluation

We carried out several experiments using our prototype device. In all experiments, both cold-wind and hot-wind conditions were reproduced.

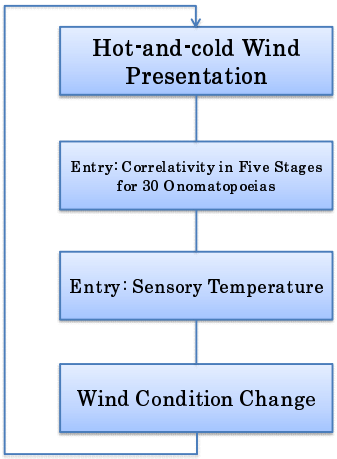


Figure 3. Flowchart of Subject Investigation

The temperature inside the temperature generator was under −30 degrees Celsius for producing cold-wind conditions and around 90 degrees Celsius for producing hot-wind conditions.

4.1 Measurement of Outlet Temperature

To establish what kinds of wind our DC fans were capable of producing, we first measured the output temperature. Placing a thermometer at the blower outlet vent, we carried out ten levels of measurement on each fan unit, starting at a PWM value of 80 (at which level the rotation of a DC fan is stable) and ending with a PWM value of 255 (the maximum output). After running a DC fan for one minute, we recorded the temperature of the thermometer. Because changes of temperature are not immediate, we ran the fan units for exactly one minute prior to measurement, and then waited exactly 5 minutes before the next operation.

Table 1. Wind Patterns Used in Experimental Investigation

Pattern	PWM Value (Propeller Fan)	PWM Value (Blower Fan)
Pattern 1	80	0
Pattern 2	160	0
Pattern 3	255	0
Pattern 4	0	80
Pattern 5	0	160
Pattern 6	0	255

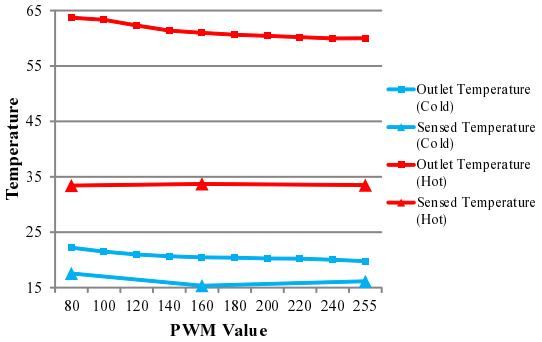


Figure 4. Comparison between Outlet Temperature and Sensed Temperature When Using Propeller Fan

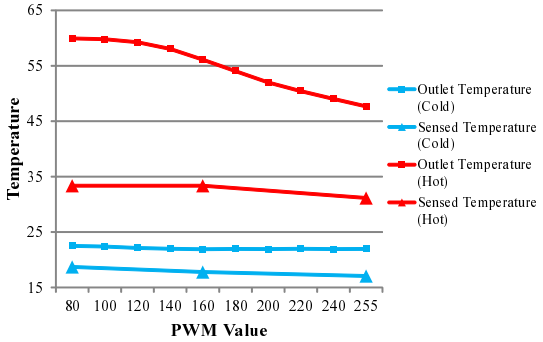


Figure 5. Comparison between Outlet Temperature and Sensed Temperature When Using Blower Fan

4.2 Investigation of Sensory Temperature

Assuming the impression a user has toward hot and cold winds will be greatly affected by the way s/he senses temperature, we sought to gauge sensory-perceptual variations among our subjects. To accomplish this, we decreased the number of wind-sensation patterns to six (three stages of PWM value for each of the fan units), so that the subjects could easily distinguish among them (see Table 1). These 6 patterns were then combined with both hot and cold sensations to produce a total of 12 test wind conditions. Fourteen male and female subjects, all in their twenties, were asked to record their sensations of these conditions.

4.3 Investigation of Correlation of Hot-and-cold Wind Conditions to Onomatopoeia

To investigate how the users correlate hot-and-cold wind conditions to onomatopoeic meanings, we presented each of the fourteen subjects with 30 onomatopoeic words related to wind and temperature, and asked them rate the degree of correlation on a five-point scale. This investigation of correlativity and the above investigation of temperature sensation were carried out at the same time. The combined flow of the two investigations is shown in Figure 3.

5 Discussion of Results

5.1 Outlet Temperature and Sensed Temperature

It is well established that the human sense of temperature is affected by changes in wind and humidity. Accordingly, we compared the outlet temperature of our blower device with the sensed temperature for both the propeller fan and the blower fan (see Figures 4 and 5).

As expected, the greater the wind strength, the lower the sensed temperature was relative to the outlet temperature. Even so, when we compared PWM value 160 to PWM value 255, the sensed temperature did not change significantly even as the outlet temperature changed greatly. We therefore conclude that wind sensation tends to blunt temperature sensation.

5.2 Discussion of Correlativity between Hot and Cold Wind Conditions and Onomatopoeia

We compared the correlativity of wind-related onomatopoeic words (e.g., pyu-, buo-) to the wind conditions produced by both the propeller fan (Figures 6 and 7) and the blower fan (Figures 8 and 9) using the same PWM values. We found that, using the blower fan, the differences in average and variance were rather small, but that, using the propeller fan, the differences in average and variance were, respectively, 2.9 and 1.5 times greater than those using the blower fan. Clearly, correlativity depended greatly on which of the two fan units was used.

For temperature-related onomatopoeic words (e.g., hinyari, pokapoka), we found that the correlativity with reproduced hot and cold temperatures was higher when using the propeller fan (Figure 10) than when using the blower fan (Figure 11). We conclude that the propeller fan unit produced stronger correlations with temperature-related onomatopoeia.

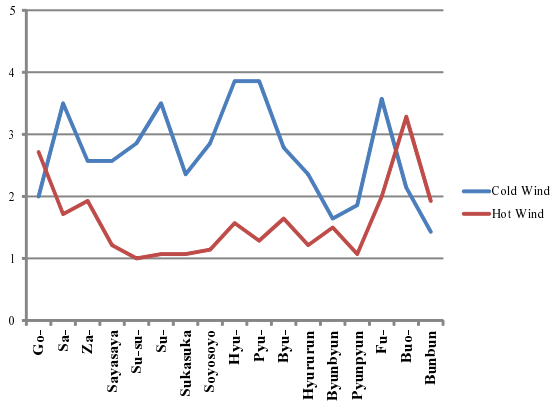


Figure 6. Average Correlativity between Wind-related Onomatopoeia and Pattern 2 Wind Conditions

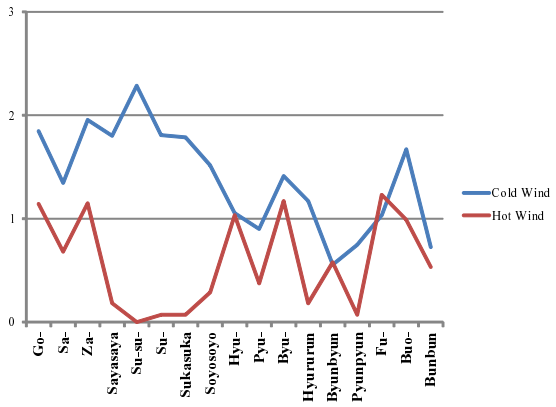


Figure 7. Variance of Correlativity between Wind-related Onomatopoeia and Pattern 2 Wind Conditions

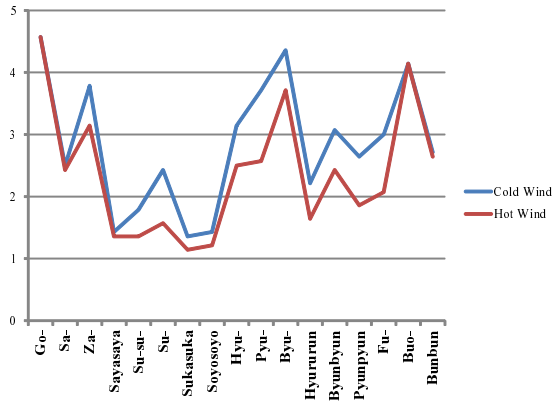


Figure 8. Average Correlativity between Wind-related Onomatopoeia and Pattern 5 Wind Conditions

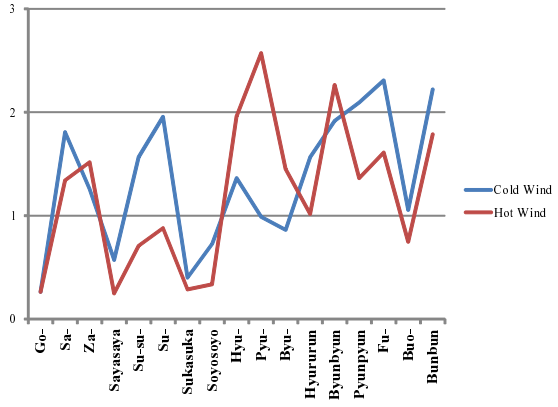


Figure 9. Variance of Correlativity between Wind-related Onomatopoeia and Pattern 5 Wind Conditions

Based on the combined results, we conclude that wind-related onomatopoeia are best conveyed using the blower fan, and temperature-related onomatopoeia are best conveyed using the propeller fan. The combination of the two fans allows us to reproduce wind and temperature conditions with high correlativity to related onomatopoeic words.

5.3 Problems and Solution

One of the challenges facing the system suggested in this paper is the inability to control temperature directly through software. Although heat can be varied by electrical current, cold cannot. In our prototype design, we had planned to accomplish temperature control using a butterfly valve attached to hot and cold air tanks. Unfortunately, problems with the valve coupling prevented proper mounting. To fix these problems, we need to experiment with valves using a different structure (e.g., gate valves).

The problem of excessive system noise, created by the fans and the PWM control, must also be addressed. The best solution may be to provide users with headphones and to present auditory information at the same time as wind information.

6 Conclusion and Future Work

In our investigation of the hot-and-cold blower device proposed in this paper, we found a fairly high correlativity between the hot and cold winds produced by our system and corresponding onomatopoeia.

In future work, we will analyze the feasibility and benefit of adding visual and auditory infor-

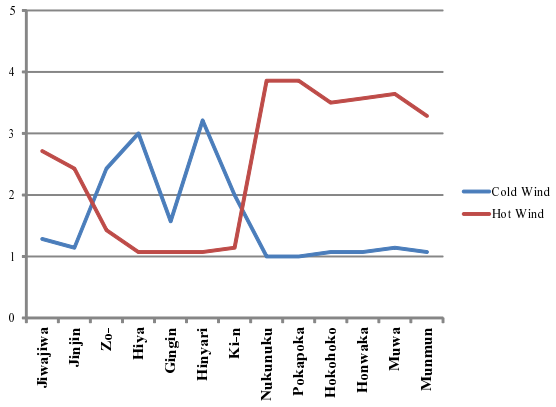


Figure 10. Average Correlativity between Temperature-related Onomatopoeia and Pattern 2 Wind Conditions

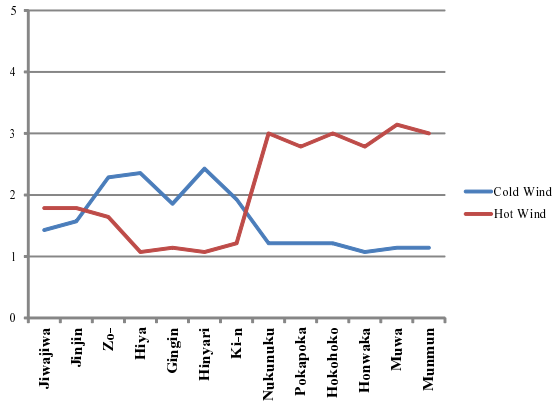


Figure 11. Average of Correlativity between Temperature-related Onomatopoeia and Pattern 5 Wind Conditions

mation to this system. We will also consider a number of presentation metaphors (e.g., intructive picture book that utilizes onomatopoeia as input[10]) that might improve the instructional value of the system.

References

[1] The Japan Foundation. *Present Condition of Overseas Japanese-Language Education Survey Report on Japanese-Language Education Abroad 2009*. The Japan Foundation, 2011.

[2] Kiyota Hashimoto and Kazuhiro Takeuchi. Prototypical Development of an E-learning System for Japanese Onomatopoeia. In *Proceedings of International Conference of Engineering Education and Educational Technology*, pages 30–33, 2010.

[3] Otodesigners Corporation. Onomatopoeia Sound Database, `http://www.otodesigners.com/sound_marketing/sound_marketing.html` (Confirmed at May 30th, 2013).

[4] Chiemi Watanabe and Satoshi Nakamura. Onomatoperori: Search Engine for Cooking Recipes by Impression of Their Tastes or Textures Using Japanese Onomatopoeia. In *Proceedings of SCIS-ISIS 2012*, PT–55, 2012.

[5] Keisuke Kambara and Koji Tsukada. Onomatopen: Painting Using Onomatopoeia. In *Proceedings of the 9th International Conference on Entertainment Computing*, pages 43–54, 2010.

[6] Takuji Narumi, Tomohiro Akagawa, Seong Young Ah, and Michitaka Hirose. Thermotaxis. *SIGGRAPH '09: Posters*, SIGGRAPH '09, Article No. 18, 2009.

[7] Kumiko Kushiyama, Momoko Inose, Rie Yokomatsu, Kinya Fujita, Toshiie Kitazawa, Mototsugu Tamura, and Shinji Sasada. Thermoesthesia: about Collaboration of an Artist and a Scientist. *ACM SIGGRAPH 2006 Sketches*, 2006.

[8] Takayuki Kosaka and Shimmi Hattori. Wind-Surround System. In *Proceedings of Interaction 2006*, pages 187–188, 2006, (in Japanese).

[9] Masahiro Ono. *Giongo Gitaigo 4500 Nihongo Onomatope Jiten (Onomatopoeia Dictionary)*. Shogakukan, 2007, (in Japanese).

[10] Hiroki Uema, Daiki Hasui, and Mitsunori Matsushita. Basic Framework on an Intructive Picture Book that Utilizes Onomatopoeia as Input. In *Proceedings of Entertainment Computing 2012*, D–09, 2012, (in Japanese).

Why are specific onomatopoeias evoked from specific images? - Investigating correlations between image features and quantified onomatopoeias

Takanori Komatsu† **Shoma Inaoka‡**

†Department of Frontier Media Science, Meiji University
tkomat@meiji.ac.jp

‡Interdisciplinary Graduate School of Science and Technology, Shinshu University
13fm204g@shinshu-u.ac.jp

Abstract

"Why are specific onomatopoeias evoked from specific images?" To tackle this question, we assumed that specific image features should be strongly related to the specific elements of quantified onomatopoeias. We then investigated the correlation between the specific image features and the specific quantified elements of user evoked onomatopoeias from such images. As a result, there was a strong correlation between the *skewness* of the image histogram and each attribute of the onomatopoeia vector for tactile-sensation-related onomatopoeias such as "SARA-SARA" and "ZARA-ZARA."

Keywords: Onomatopoeias, Sound Symbolism, Image Histogram

1 Introduction

Onomatopoeias - echoic, imitative, or mimetic words - are indispensable parts of the Japanese language [1,2]. In particular, these onomatopoeias are quite good for describing certain phenomena or events that cannot be described well literally. For example, most Japanese use the onomatopoeia "MOFU-MOFU" when they are asked to describe an image of dogs with long-hair coats like that shown on the left side of Figure 1, and they also use "KASA-KASA" when they describe the image of dry falling leaves on the right side of Figure 1. *"Why are specific onomatopoeias evoked from specific images?"* To tackle this question, we assumed that specific image features should be strongly related to the specific elements of onomatopoeias.

Fig. 1. Dogs with long-hair coats (left) and dry falling leaves (right)

In general, most Japanese onomatopoeias can be explained by the concept of sound symbolism [3,4,5], which is the idea that phonemes carry meaning in and of themselves regardless of the language, e.g., consonant /s/ has a feeling of light friction, smoothness or speed, /k/ and /t/ are hard or stiff, and /n/ has a feeling of stickiness. Therefore, it seems that the concept of sound symbolism has a strong correlation with image features like those shown in Figure 1.

The purpose of this study is then to investigate the correlation between specific image features and the specific elements of the user evoked onomatopoeias from such images. We focused on the image histogram for the specific image features and on the three-dimensional onomatopoeia vector {*sharpness, softness, dynamics*} [6] for the specific elements of the evoked onomatopoeias. We then conducted an experiment to make participants answer with the appropriate onomatopoeias to describe the presented images and analyzed the correlation between the characteristics of image histograms and each attribute value of the three-dimensional onomatopoeia vector.

2 Experiment

2.1 Images

We prepared the 15 different images shown in Figure 2. These images were taken from sections of whole pictures of daily clothing (e.g., denim/down/leather/boa jacket), and these were taken in the same lighting condition. The size of these images was 3,264 × 1,836 [pixel] (about 3.5 × 2.0 [cm] on a 14-inch laptop PC display), and these images were gray-scaled in 8-bit resolution.

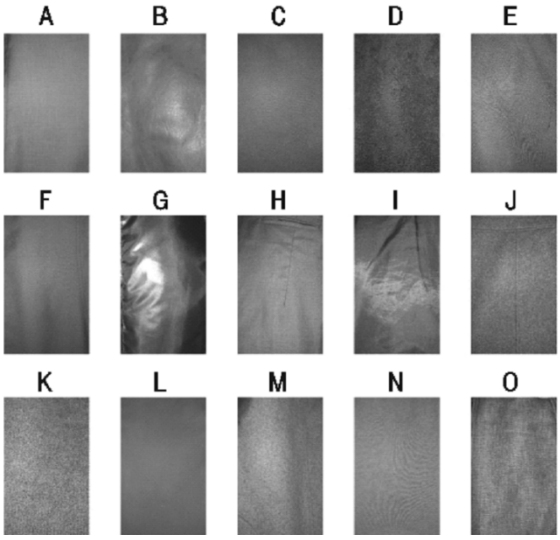


Fig. 2. 15 different pictures

2.2 Image features

As specific image features, we focused on the image histogram, which plots the number of pixels for each tonal value (Figure 3). Therefore, this histogram can act as a graphical representation of the tonal distribution in a digital image. In fact, the image histogram is widely used for comprehending the characteristics of images and is already applied to most digital cameras on the market. In addition, Motoyoshi et al. [7] reported that human perception of the glossiness and lightness of various natural surfaces is well correlated with skewness in an image histogram. Therefore, we chose the image histogram for the specific image features in this study.

In particular, we focused on four specific features of the image histogram: *average, median, skewness, and kurtosis*. These four values were calculated by using ImageJ ver. 1.44b[1], a public domain, Java-based image processing program. Table 1 shows the average, median, skewness, and kurtosis of the image histogram for each image in Figure 2.

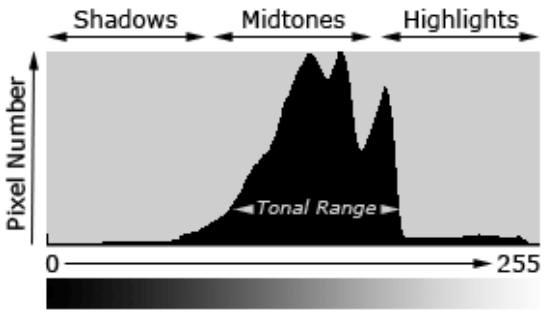


Fig. 3. Example of image histogram

Table 1. Image features for each image

Image	Material	Average	Median	Skewness	Kurtosis
A	tailored jacket	129.97	131	-0.7	1.76
B	leather jacket	121.34	118	1.37	3.69
C	boa jacket	102.7	103	0.11	-0.3
D	boa pile-woven	75.68	75	0.35	0.33
E	gilet	115.89	113	0.46	-0.14
F	tailored jacket	102.81	103	0.04	-0.11
G	down jacket	98.58	94	0.71	0.38
H	gilet	125.26	128	-0.59	0.39
I	gilet	109.07	106	0.74	1.1
J	pea coat	110.28	108	0.52	0.27
K	cardigan	127.31	128	-0.45	0.58
L	tailored jacket	99.1	99	0.12	-0.46
M	sweatshirts	129.05	129	0.1	-0.33
N	cotton jacket	126.22	124	0.21	-0.005
O	denim jacket	118.28	114	0.61	0.28

2.3 Elements of onomatopoeias

As elements of onomatopoeias evoked from the images, we focused on a three-dimensional onomatopoeia vector proposed by Komatsu [6]. Specifically, Komatsu proposed the "onomatopoeia vector" in order to quantify XYXY-type onomatopoeias (i.e., onomatopoeias consisting of a repetition of two syllables, X and Y) that have the three attributes {*sharpness, softness, dynamics*} on the basis of the concept of sound symbolism. Each attribute can be calculated in the following equations in (1).

$$\begin{cases} I_{\text{sharpness}} = 0.60S_1 + 0.52S_2 \\ I_{\text{softness}} = 0.56S_1 + 0.46S_2 + 0.22B_2 \\ I_{\text{dynamic}} = 0.59S_1 + 0.44S_2 \end{cases} \quad (1)$$

Here, S_1, S_2, and B_2 are the values of the consonant vectors of X and Y and the vowel vector of Y, respectively. Komatsu [6] already prepared the consonant and vowel vectors in order to express onomatopoeia vectors by combining these consonant and vowel vectors. The most significant advantage of this method is that any XYXY-type onomatopoeia can be quantified even if it is not listed in any dictionary.

[1] http://rsbweb.nih.gov/ij/

In particular, we focused on each of the three attributes of the onomatopoeia vector independently.

2.4 Procedure

In this experiment, 30 university students (15 men and 15 women; 19 - 25 years old) participated. First, the participants were asked to take a seat in front of laptop PC. Figure 2 then appeared on the PC's display, and they were asked to answer with only one appropriate onomatopoeia on a response sheet for each image (A - O).

2.5 Results

Acquired onomatopoeias: On average, 12.26 different kinds of onomatopoeias were listed for each image (minimum: 7 onomatopoeias, maximum: 16). The average rate for how much the mode onomatopoeias (most frequently listed onomatopoeias) appeared among all 30 answers in each image was 40.2%, so we determined that each mode onomatopoeia in each image were a representative onomatopoeia of the image. Table 2 shows the representative onomatopoeias and their onomatopoeia vectors.

Table 2. Representative onomatopoeias for each image

Representive Onomatopoeias	Material	Onomatopoeia vector		
		Sharpness	Softness	Dynamics
SARA-SARA	A, C, F, H, K	0.704	1.0412	1.0233
ZARA-ZARA	D, E, J, M, N, O	0.662	0.1676	0.864
TSURU-TSURU	B, I	0.422	0.8076	1.0115
TEKA-TEKA	G, L	1.436	-0.5496	1.3943

Table 3. Correlation matrix for all 15 images

	sharpness	softness	dynamics	average	median	skewness	kurtosis
sharpness	-	*-0.683 ***	*0.837 ***	-0.348 ns	-0.342 ns	-0.106 ns	-0.402 ns
softness		-	*-0.349 ns*	0.369 ns	0.428 ns	-0.381 ns	0.366 ns
dynamics			-	-0.277 ns	-0.267 ns	-0.014 ns	-0.013 ns
average				-	*0.991 ***	*-0.305 ns*	*0.264 ns*
median					-	*-0.413 ns*	*0.226 ns*
skewness						-	*0.356 ns*
kurtosis							-

Correlation between image features and onomatopoeia vector for all 15 images: To investigate the correlation between the image feature and the elements of acquired representative onomatopoeias, the correlation coefficients between 4 image features (average, median, skewness, and kurtosis) and 3 attributes of representative onomatopoeias (sharpness, softness, and dynamics) were calculated by using 15 images. Table 3 shows the correlation matrix among these four 4 image features of the 15 images and the 3 attributes of the corresponding onomatopoeias. As the result of this analysis, it was found that there were no significant correlations between the image features and the elements of the onomatopoeias.

Rethinking onomatopoeias: We then re-thought the acquired four different kinds of onomatopoeias. In general, "SARA-SARA" is used for expressing something smooth, light, or softly flowing, and "ZARA-ZARA" is for a rough and coarse feeling, so it is said that these two onomatopoeias are tactile-sensation-related onomatopoeias, while "TSURU-TSURU" is used for expressing something shiny or slippery, and "TEKA-TEKA" is for something shiny or bright, so it is said that these two onomatopoeias are visual-sensation-related onomatopoeias.

Some studies already reported that most Japanese onomatopoeias are related to tactile sensations [8,9], and in our experiment, 11 out of the 15 images evoked the tactile-sensation-related onomatopoeias "SARA-SARA" and "ZARA-ZARA." Also, the 15 images in this experiment were of daily clothing, so it would be reasonable for the participants to answer with tactile-sensation-related onomatopoeias. Therefore, we re-analyzed the correlation coefficients by using 11 images for which the representative onomatopoeias were "SARA-SARA" or "ZARA-ZARA."

Correlation between image features and onomatopoeia vectors for 11 images: The correlation coefficients between 4 image features and 3 attributes of representative onomatopoeias were calculated by using 11 images. Table 4 shows the correlation matrix among these 4 image features of the 11 images and 3 attributes of the corresponding onomatopoeias.

Table 4. Correlation matrix for 11 images

	sharpness	softness	dynamics	average	median	skewness	kurtosis
sharpness	-	*1.000 ***	*1.000 ***	0.161 ns	0.253 ns	**-0.800 ****	0.355 ns
softness		-	*1.000 ***	0.161 ns	0.253 ns	**-0.800 ****	0.355 ns
dynamics			-	0.161 ns	0.253 ns	**-0.800 ****	0.355 ns
average				-	*0.993 ***	*-0.457 ns*	*0.259 ns*
median					-	*-0.553 +*	*0.288 ns*
skewness						-	*-0.605 **
kurtosis							-

As the result of this analysis, it was found that there were significant strong negative correlations between the three attributes of the onomatopoeia vector {*sharpness, softness, dynamics*} and *skewness* in the image histogram, respectively ($r = -0.8, F = 16.02, df1 = 1, df2 = 9, p$

< *.01*). Therefore, the results showed that there was a significant difference for the image feature "skewness" between the images that evoked the tactile-sensation-related onomatopoeias "SARA-SARA" and "ZARA-ZARA."

Figure 4 shows the actual image histograms that evoked "SARA-SARA" for the participants (images A, H, and K in the top half of Figure 4) and those that did so for "ZARA-ZARA" (images B, N, and O on the lower half of Figure 4). Figure 4 clearly shows that the peaks of the histograms in the upper images (A, H, and K) slightly shifted to the right, while those in the lower figures (B, N, and O) shifted to the left.

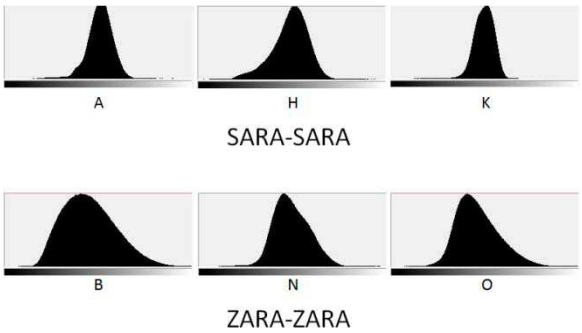


Fig. 4. Typical image histograms that evoke "SARA-SARA" (upper) and "ZARA-ZARA" (lower)

3 Discussion and conclusions

The results of this experiment clarified the correlation between the specific image features and the specific elements of the quantified onomatopoeias; that is, the *skewness* of the image histogram strongly correlated with each attribute of the onomatopoeia vector {*sharpness, softness, dynamics*} in the case of tactile-sensation-related onomatopoeias such as "SARA-SARA" and "ZARA-ZARA."

However, the variations of tactile-sensation-related onomatopoeias in this study were limited to only two, so it is definitely necessary to include a much wider variety of tactile-sensation-related onomatopoeias in our future study.

In this analysis, we ultimately eliminated the visual-sensation-related onomatopoeias "TSURU-TSURU" and "TEKA-TEKA"; however, the image histogram of "TEKA-TEKA" was quite easy to distinguish from the other kinds of onomatopoeias (Figure 5). Therefore, we should focus on the other kinds of image features, not only the four features used in this study but also others such as the number of local peaks, the values of the integrals, and so on in order to much more comprehend the general tendency of the correlations between image features and onomatopoeias.

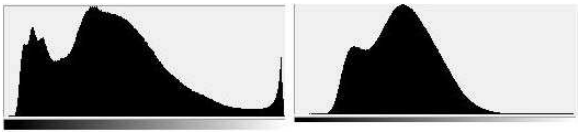


Fig. 5. Image histogram that evoked "TEKA-TEKA"

Acknowledgment

This study was supported by JSPS KAKENHI Grant Number 25330319.

References

[1] Tamori, I. 1988. Japanese onomatopoeias and verb-less expressions, Jimbun Ronshu, 24, 2, 1-25.

[2] Millington, S. 1993. NIHONGO PERA PERA!: A User's Guide to Japanese Onomatopoeia, Charles E. Tuttle Company, Rutland, VT.

[3] Sapir, E. 1929. A study of phonetic symbolism. Journal of Experimental Psychology. 12, 4, 225-239.

[4] Köhler, W. 1929. Gestalt Psychology. Liveright, NY.

[5] Ramachandran, V. S., and Hubbard, E. M. 2001. Synaesthesia: A window into perception, thought and language. Journal of Consciousness Studies, 8, 12, 3-34.

[6] Komatsu, T. 2012. Quantifying Japanese Onomatopoeias: Toward Augmenting Creative Activities with Onomatopoeias, In Proc. of AH12, a-15.

[7] Motoyoshi, I., Nishida, S., Sharan, L. and Adelson, E. H. 2007. Image statistics and the perception of surface qualities. Nature, 447, 206-209.

[8] Watanabe, J., Utsunomiya, Y., Tsukurimichi, H., and Sakamoto, M. 2012. Relationship between Phonemes and Tactile-emotional Evaluations in Japanese Sound Symbolic Words, In Proc. of CogSci2012, 2517-2522.

[9] Watanabe, J., and Sakamoto, M. 2012. Comparison between onomatopoeias and adjectives for evaluating tactile sensations, In Proc. SCIS-ISIS2012, 2346-2348.

Constructing a Knowledge Base of Onomatopoeia-Action Relations for an Interactive Illustrated Storybook

Hiroki Uema† **Daiki Hashui**† **Mitsunori Matsushita**‡
†Graduate School of Informatics, Kansai University
‡Faculty of Informatics, Kansai University
mat@res.kutc.kansai-u.ac.jp

Abstract

This paper presents an interactive illustrated storybook for PCs and tablets, which can impart virtual actions to the characters and phenomena (e.g., "raining," "flowers swaying") in a story on the basis of onomatopoeias (e.g., "Para Para," "Byu- Byu-.") input by the user. In order to create a knowledge base for this system, we investigate the correspondence between the onomatopoeias and the characters/phenomena using tweets containing onomatopoeias. This paper describes the proposed system and the method for obtaining the above mentioned correspondence.

Keywords: Onomatopeia, Electric Storybook

1 Introduction

Smartphone and tablet devices have become increasingly popular in recent years. These devices make e-books increasingly usable and popular. Many types of e-books including novels, guidebooks, and comics have gone into circulation. Owing to the digitization of book content, moving a picture or characters with sounds on e-book devices is now possible. In addition, such content can be altered interactively via touch-sensible panels.

However, many e-books still contain only simple graphics, and the number of e-books with interactivity is quite low. Two reasons can be conjectured: one is that representation techniques suitable for e-books are not well established, and the other is the lack of tools for creating such representations smoothly. These problems pose difficulties for creators when they produce dynamic content [1].

As the first step to solve the problems, we are attempting to clarify a set of primitives for creating dynamic content. Some dynamic content, such as animated film (e.g., movies or animations) and videogames, is now popularly used on e-book devices. Although one could argue that animated films effectively exploit the advantages of e-books, this type of content is passive and lacks interactivity because users only view information presented by devices. In contrast, a substantial portion of video game content involves active content (i.e., their users communicate the content interactively). We believe that content producers can make use of the advantages of e-books by adding in content not only with dynamic elements but also participating elements such as video games.

On the basis of this perspective, the primary aim of the present study was to create interactive e-book content that involves user participation, such as the presentation of dynamic content by user key in. In this study, which is the first of several ongoing projects, we describe an onomatopoeia picture book system. The system provides interactive e-book content that can attach dynamic expressions to objects in the system by means of input onomatopoeias. In addition, we state how to create a knowledge base for this system, which contains the correspondence between the onomatopoeias and objects/phenomena. We provide a method to construct this knowledge base by collecting tweets that contain onomatopoeia via Twitter and to check the correspondence between the onomatopoeias and objects/phenomena.

2 Related Works

We propose the use of onomatopoeia to consolidate a new and easy-to-use method of expressing e-book content. Recently, onomatopoeia, which is a method of expressing human ideas relating to an object by the sounds with which it is associated, has attracted attention cross-sartorially. For example, it is studied in engineering as a potential method to express human ideas by auditory input provided to a computer.

"Moyamoya Drawing," which was proposed

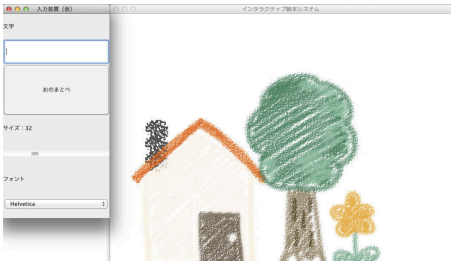


Figure 1. Onomatopoeia storybook system PC version

by Terashima, uses intuitive impressions in the manner of onomatopoeia [2]. The system uses an impression to select an effect for editing an image. This is intended to help beginners who are unaware of the details of the effect when they edit the image.

"Onomatoperori" allows the user to represent the impression of a meal by using onomatopoeia. The system helps the user by recommending an appropriate recipe [3]. Onomatopoeic words often used to represent the texture of the food (e.g., "fuwa fuwa," "kari kari") in Japanese are used to search for recipes.

Allowing the use of onomatopoeia to represent a meal that the user wants to cook improves user convenience. In this study, we use onomatopoeia to add movement to a picture in a dynamic e-book. Our proposed feature will help to generate dynamic representations in accord with the intuition of the user.

3 Packaging of Onomatopoeia Storybook System

The onomatopoeia picture book system that we propose changes the situation or moves objects (e.g., "House," "Tree," "Flower") when an onomatopoeic word is entered.

The picture book system is implemented as follows, a house, trees, people, and flowers are among the objects displayed in the picture book (See Figure 2). Changes of weather (e.g., rainy, windy, sunny) are also displayed. Onomatopoeic words associated with these objects and conditions were examined from the onomatopoeia-related term knowledge base that is prepared for the picture book system. More information on this knowledge base will be described later.

Figure 2. Onomatopoeia storybook system tablet version

4 System Structure

In the current implementation, we have produced two picture book systems that run on an Android tablet or a PC. The user can input an onomatopoeic word by using a button arranged on the screen in the Android version or a drop-down list and text forms in the PC version.

The PC version consists of two windows: an interface window to enter onomatopoeic words, and a canvas window on which a picture is drawn by the system (See Figure 1). In the text form of the window for onomatopoeic input on the PC version, the "motion" or "status" of a picture displayed in the canvas window can be changed by filling in the 22 onomatopoeic words on the system and clicking the "onomatopoeia" button. Another set of 22 onomatopoeic words that can be applied to the system is shown in the drop-down list. The user can select a word from the list and enter it as input by clicking the "onomatopoeia" button.

The Android version is divided into upper and lower screens (See Figure 2); the upper canvas displays a picture, and the lower canvas allows the provision oqf a button that can be used to enter information about onomatopoeia. Furthermore, the picture can be moved by pressing the button indicated by the word onomatopoeia. Ten buttons are displayed on the screen. This allows button sizes that can be pressed simultaneously to display the corresponding objects (e.g., flower, person).

Both Android and PC versions of the above system have been developed. Each object in

the picture drawn is a separate layer, and multiple objects are displayed through overlaying. There is no difference in operation between the Android and PC versions (for entering onomatopoeia, to impart a dynamic representation corresponding to the input onomatopoeia for the associated objects).

5 Activity of System

In the two systems described above, some objects are placed on the canvas part in the initial state. For example, a painting of a "house, tree, and flower" is displayed on the canvas part in Figure 1. By entering an onomatopoeia associated with the picture on the screen, a user can cause a change (e.g., move the object that corresponds with the onomatopoeia or change the weather) in the picture book. For example, Figure 3 is an example of a change that occurs when the user enters the onomatopoeia of "za- za-." In the onomatopoeia picture book system, the weather in the default state is sunny. This can be changed to rainy weather by entering the weather-related onomatopoeia of "za- za-" therein.

The change in the behavior of the object to be drawn is determined on the basis of the correspondence between the onomatopoeias and the objects/phenomena. For example, when an onomatopoeia called "Moku moku" is entered, the system displays a situation wherein the object that is attached corresponds to the "smoke" drifting up a chimney. If a visual representation is difficult, such a system is expressed using the motion of objects that are not directly related to the input onomatopoeia. For example, the onomatopoeia of "byu-" corresponds to a change in the conditions of "wind" although "wind" cannot be viewed. Therefore, such effects are shown through the shaking of "flowers," the direction in which something falls, like the falling of "leaves" or "rain" (See Figure 4). In the current implementation, the mapping between the onomatopoeia and the behavior of each object is performed manually in advance. For the PC version, we were able to enter the onomatopoeia of 22 types of objects. Some examples of operations and associated onomatopoeia are shown in Table 1.

The intensity of an operation can also be adjusted by using different onomatopoeia. For example, both "Shito shito" and "za- za-" relate to

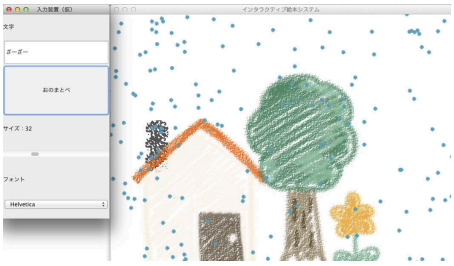


Figure 3. If user enter "za- za-"

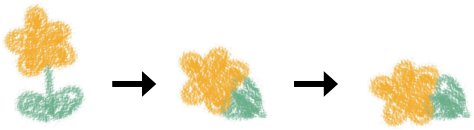


Figure 4. If user enter "pyu-"

"rain." However, there is a difference between the two; in this case, one produces a dull sound whereas the other does not. Here, sound symbolism of the onomatopoeia of these two types of rain can be used to differentiate between the degree of "rain." There is systematic relationship between sound and meaning as stated by Ohara [4]. By referring to this study, we correlated "heavy rain" with "za- za-" and "sprinkling rain" with "Shito shito." In this system, onomatopoeic text can be entered only one at a time. Therefore, to display simultaneous events, the user can input the order in which the onomatopoeias must occur. For example, wind and rain can be superimposed to occur simultaneously by entering "hyu-" immediately after za- za-. However, entering San-san (sunny) when raining is incompatible. Therefore, if the onomatopoeia above is entered as input; the events occur in sequence, i.e., "the sun comes out after the rain stops."

Table 1. Correspondence of system operation and onomatopoeia

Input onomatopoeia	System operation
gasa gasa	Falling leafs(from tree)
gata gata	Bouncing a house
gira gira	The sun be vivacity/Be clear
san san	The sun be vivacity/Be clear
sa-	Rain(Weak)
za- za-	Rain(Gusty)
:	:

6 Onomatopoeia Related Term Knowledge Base

The onomatopoeia picture book system we propose in this paper allows a picture to be changed through onomatopoeic input. To achieve this, each onomatopoeic word must be clearly defined to correspond to either a "motion" or a "changing circumstance."

For example, the onomatopoeia for "uro uro" can mean both "walking" and "running" [5].

However, to find the onomatopoeia for "uro uro" associated with the word for "walk," or conversely, to find the words corresponding to the onomatopoeia "uro uro," certain associations must be made.

One way to do this is to obtain peoples' opinions through a questionnaire. However, such an approach would take time and is impractical.

Thus, we wish to find an alternative, automatic approach to finding relevant words corresponding to onomatopoeia and vice-versa.

6.1 Research Method of Onomatopoeia

The representation content of onomatopoeia can be examined using a Japanese onomatopoeia dictionary [5]. This dictionary contains around 4500 words, which are arranged by "sound" and "state" and the definitions indicate whether the onomatopoeia are based on sound or appearance. However, it is not sufficiently clear whether an onomatopoeic word corresponds to a situation or a movement of an object. In addition, identifying a newly created onomatopoeia is impossible.

In our opinion, there are about 4500 Japanese onomatopoeic words, and a corpus containing numerous onomatopoeic words is required to collect all related terms. For this purpose, a Web-based text corpus is preferable.

On the basis of the above discussion, a knowledge base was constructed to examine the relationship between words and onomatopoeia so as to prerecord related words corresponding to the onomatopoeia and the onomatopoeia itself. However, manually matching all 4500 Japanese onomatopoeia is a tedious task. Thus, we investigated the automatic construction of the knowledge base containing onomatopoeia-related terms.

We have speculated that this knowledge base may be used to determine the relationship between onomatopoeia and words by examining sentences that contain onomatopoeia related to a person performing informal communication, because we have focused on the use of onomatopoeia when a person speaks. Specifically, we collected the text that contains the onomatopoeia by using the Twitter mini blog service, and we checked for the co-occurrence frequency of words that occur together with the onomatopoeia included in the text. Further, we examined whether high co-occurrence frequency is related with onomatopoeia. Then, by creating a knowledge base, we developed mechanisms for examining the correspondence between related words and onomatopoeia.

6.2 Collect Tweets Containing Onomatopoeia

Because the proposed method is intended to be used for movements drawn in illustrated books, onomatopoeia are used in its implementation. We collected onomatopoeia that can express the movement of human. Onomatopoeia-related words that are basic movements of a person, e.g., "aruku (= walking)" and "hashiru (= running)," were used to create an onomatopoeia meter — 56 words that are described in the Japanese Onomatopoeia Dictionary.

Subsequently, morphological analysis was carried out using MeCab [6]. The collected tweets were divided into nouns, verbs, and adjectives, and the co-occurrence frequency of each word was determined.

In the case of part-of-speech determination by MeCab, symbols such as "@" were identified as nouns. In this study, we measured the co-occurrence frequency of related Japanese words and Japanese onomatopoeia.

Table 2 shows the co-occurrence frequency obtained using the onomatopoeia "bra bra."

7 Analysis of Related Terms Knowledge Base

As mentioned previously, we searched for onomatopoeia words related to "walk" to "run" using Twitter API 56 and collected sentences that contains the onomatopoeia. The collection of tweets was performed over 10 days, from December 11, 2012 to December 21, 2012, and we collected 265650 tweets.

We examined the degree of co-occurrence for the verbs "walk" and "run," and each onomatopoeic word for the word co-occurrence fre-

Table 2. Noun words co-occur with onomatopoeia “bura bura”

Co-occur noun word	Co-occurrence frequency
(n)	1146
Ł(kyou)	1108
(no)	1043
(warai)	716
(asa)	677
(sann)	668
(hi)	599
(koto)	484
Ł(ima)	619
(mati)	582
(eki)	579
:	:

quency of tweets contains the onomatopoeia as well as related terms from the knowledge base.

Because the number of tweets collected was less than the number of collected onomatopoeia, the extra words were excluded from the analysis of 15 words (e.g., “pappaka,” “shanari shanari”) out of 56 collected onomatopoeic words. Moreover, it was confirmed that the maximum contents are of tweets containing “ta-,” the most tweeted phrase. For onomatopoeia such as “yokatta-” and “yatta-,” we could not collect tweets containing “ta-.” Thus, we excluded them from the analysis of the collected results. We analyzed a total of 40 words, excluding the 16 collected from 56 onomatopoeias.

In the analysis, we examined whether the results included the words “walk” or “run” to identify co-occurrences and the correspondence degree of co-occurrence frequency.

Table 3 shows the co-occurrence rank of “walk” and “run” in terms of the number of co-occurrences of each onomatopoeic word. On examining the co-occurrence frequency, it was found that “walking” or “running” co-occur with 39 onomatopoeic words out of 40. It was ranked among the top 10 with regard to co-occurrence verbs, accounting for 29 out of 40 onomatopoeic words.

These results confirm that our approach obtains co-occurrence related (e.g., “walk,” “run”) and onomatopoeic words.

However, only the onomatopoeia “taji taji” did not co-occur with “walking” or “running” (See Table 3). This is presumably because it is often used in the sense of “appearance of falter” and “appearance of a people have trouble,” but is not commonly used in the sense of walking.

Table 3. Number of co-occurrences between onomatopoeias and verbs

Onomatopoeia	**Walk**	**Run**	**# of verbs**
tyoko tyoko	65	97	3291
hura hura	14	71	3125
uro uro	40	137	2619
bura bura	19	180	2138
bata bata	260	61	1976
sassa	79	221	1583
tototo	256	195	1332
yoro yoro	20	79	1267
noronoro	9	12	1241
teku teku	2	77	1184
to-n	293	672	1160
noko noko	20	221	1091
yoti yoti	4	104	1059
dota bata	139	58	1035
wata wata	587	98	1042
yobo yobo	16	80	1028
tobo tobo	1	36	1002
noso noso	7	56	960
bata bata	74	23	862
toko toko	3	19	819
pappa	363	267	747
tatatata	79	91	600
yota yota	3	26	593
hyoko hyoko	4	24	567
suta suta	2	49	549
:	:		
seka seka	11	79	299
dadadada	-	135	280
taji taji	-	-	266
rattatta	121	66	256
sutakorasassa	97	131	250
nosi nosi	9	30	220
nossi nossi	2	65	139
syanari syanari	2	-	104

8 Conclusion

In this study, we focused on “onomatopoeia” as a means for performing on the basis of the human sense. We proposed an interactive picture book system that uses onomatopoeia as the input to operate dynamic content with interactivity of user participation. In the proposed system, the painting system can be modified by entering onomatopoeic expressions that are sensuous to the user.

To realize such a system that uses ono-

matopoeia, the ways in which onomatopoeic words and their contents should correspond must be examined. Therefore, we proposed a method to determine the co-occurrence frequency of related words and onomatopoeic words by using sentences that contain onomatopoeic words collected from Twitter.

In the future, we will improve the interface of the onomatopoeia picture book system, such that it constantly evolves, e.g., corresponding to other onomatopoeia. While striving to achieve input methods based more on the intuition of the user, we will continue to conduct research with the aim of improving the accuracy of the onomatopoeia-related word knowledge base.

References

[1] Mitsunori Matsushita and Natsumi Imaoka. Kinetic onomatopoeia generation system for creating an attractive digital comic. In *Proceedings of the 25th Annual Conference of the Japanese Society for Arti cial Intelligence*, number 1C1-OS4a-3, pages 1–4, 2011.

[2] Hiroki Terashima and Takanori Komatsu. Moya-moya drawing: Proposing a drawing tool system which can utilize users' expressed onomatopoeias as an drawing effect. In *Proceedings of the 6th International Conference on Soft Computing and Intelligent Systems and the 13th International Symposium on Advanced Intelligent Systems (SCIS-ISIS2012)*, pages 494–497, 2012.

[3] Kanwipa Lertsumruaypun, Chiemi Watanabe, and Satoshi Nakamura. Onomatoperori: Search engine for cooking recipes by impression of their tastes or textures using japanese onomatopoeia. In *Proceedings of the 6th International Conference on Soft Computing and Intelligent Systems and the 13th International Symposium on Advanced Intelligent Systems (SCIS-ISIS2012)*, pages 2349–2350, 2012.

[4] John J. Ohala. Sound symbolism. In *Proceedings of the 4th Seoul International Conference on Linguistics [SICOL].*, pages 1–6, 1997.

[5] Masahiro Ono. *Giongo Gitaigo 4500 Nihongo Onomatope Jiten (Onomatopoeia Dictionary).* Shogakukan, 2007.

[6] Taku Kudo. Mecab : Yet another part-of-speech and morphological analyzer. *http://mecab.sourceforge.net/*, 2005.

Comparison of onomatopoeia use among areas using Diet records of Japan

Sachiko Hirata-Mogi† **Satoshi Nakamura‡** **Takanori Komatsu‡** **Kimi Akita***

†School of Medicine, the University of Tokyo, Japan
‡Department of Frontier Media Science, Meiji University, Japan
*Graduate School of Language and Culture, Osaka University, Japan
marshmallow1214@gmail.com

Abstract

Onomatopoeia is believed to be particularly frequent in some regions of Japan. In this study, we examined whether there is really a difference in the use of onomatopoeia among the eleven major regions of Japan. For this purpose, we used conversation data from the most recent 2 decades of Japanese Diet Records, along with the hometowns of the speakers. The results suggested that there are no differences in frequency of use across all types of onomatopoeia in each region. However, one type of onomatopoeia, especially emphatic onomatopoeia, showed regional differences. Further research from a dialectic, micro-scale perspective is needed.

Keywords: Onomatopoeia, Dialect, Region Comparison, Diet Records, Corpus Study

1 Introduction

Japanese onomatopoeic words (also called mimetics, sound-symbolic words, idiophones and iconic words) are more frequent in spoken discourse, especially in informal daily conversations, than in writing. It is a common belief that onomatopoeia is particularly frequent in some areas, such as the Kinki region, which includes Osaka, Kyoto, Nara, Hyogo, Shiga, and Wakayama. To examine the plausibility of this folk dialectology, we investigated the frequency of onomatopoeia in the Minutes of the Diet as a corpus of spoken Japanese.

1.1 Onomatopoeia

Onomatopoeic words imitate sounds (or even soundless situations) or personal feelings. There are several categories of Japanese onomatopoeia; for example, *gion-go* (e.g., “don-don” as the sound of drum) imitates sounds with voices, and *gitai-go* (e.g., “doki-doki” as being nervous) imitates the situation or personal feelings.
Onomatopoeia forms an important part of the Japanese vocabulary and is widely used in product advertising and in comics. Appropriate contexts and situations for the use of onomatopoeia have been described. For example, previous research has shown that it is not commonly used with older persons or with superiors in a business situation [4]. On the other hand, it is used in conversations with older and younger people in personal relationships, including relatives. In this way, the number of people with whom it is appropriate to use onomatopoeia in daily conversation is relatively limited.

1.2 Onomatopoeia and region

Most Japanese people tend to believe that people who live in the Kinki region have a higher frequency of onomatopoeia use than those in other regions. Descriptions of onomatopoeia can be found in studies of the Osaka/Kansai dialect and Osaka culture. But contrary evidence comes from a dialect investigation, which reported that the amount of onomatopoeia included in conversations in Kinki region was not so great, and the amount was relatively greater in Tohoku and the Kanto region [2].
In a preliminary study, we conducted a web questionnaire to reveal the subjective frequency

of onomatopoeia use in Japan. In this study, 15- to 65-year-old men and women who lived in 47 prefectural and city governmental districts in Japan (total 1100 people) were asked how they usually use onomatopoeia [3]. There were no significant differences in subjective frequencies of onomatopoeia use among the 11 regions of Japan. However, to the question ”People in which region do you think use more onomatopoeia?”, most respondents reported the people in Kinki region These results indicated that the folk theory -- "people who live in the Kinki region have a higher frequency of onomatopoeia use than those who live in other regions" -- did not match the facts. We concluded that this bias came from the stereotype of people in the Kinki region provided by the mass media. A limitation of this study is that the subjective data may not have exactly reflected the frequency of individual onomatopoeia use because subjective reports of frequency are not always in accord with real conversation. Moreover, the respondents to a web questionnaire may not have had an adequate understanding of onomatopoeia.

1.3 Purpose of this study

To resolve the subjectivity of frequencies of onomatopoeia use in the previous study [3], we selected the Diet Records of Japan as a massive conversation corpus. Diet Records of Japan is the largest corpus of conversation we could obtain. Moreover, almost all speakers in this record are public people and it is relatively easy to trace their hometown. From these reasons, we selected this record as most appropriate data to demonstrate about the onomatopoeia usage. The purpose of our research was to reveal differences in rates of onomatopoeia usage among regions in Japan based on the Diet Records. For this purpose, we analyzed the onomatopoeia usage rates and the hometowns of the most frequent users of onomatopoeia based on the Diet Records.

2 Methods

2.1 Diet Records of Japan

The Diet Records of Japan include all utterances from all meetings of the Diet held from 1947 to 2011, with accompanying information like speaker names and dates of the meetings. This database is open to the public on the Diet Records Search System website (http://kokkai.ndl.go.jp/). It contains the meeting records of the Houses of Representatives and Councilors, plenary sessions of both Houses, and meetings of the conference committee and the joint examining committee. To compare with our preliminary study in 2011, we limited the data to Houses of Representatives and Councilors plenary sessions from 1992 through 2012 to avoid historical influence of old conversations.

2.2 Onomatopoeia list

We used an onomatopoeia list created from a onomatopoeia dictionary [1]. This list had seven categories of onomatopoeia sorted by morphological characteristics (see Table 1). We collected all the utterances included at least one form of onomatopoeia from the corpus. We calculated the frequency of onomatopoeia use (the number of utterances of each individual which included onomatopoeia divided by his/her total number of utterances). To prevent the extraction of non-onomatopoeia words, we added “-to” to the end of each search word, using the characteristic of onomatopoeia that it frequently used with Japanese suffix “-to”.

2.3 Analysis

Each data record from the Diet Records of Japan consisted of the date and time of the utterance, the name and position of the speaker, and the utterance itself. The data were arranged in chronological order from earliest to latest.
First we extracted the date information of from each meeting and converted it to western calendar notation. Then we extracted information about the speakers, their positions, and their utterances. The utterances were broken into sentences, and the database therefore consisted of individual sentences with IDs to identify unique utterances. A sentence was defined as the portion of an utterance between two periods or the point where the speaker changed. The number of sentences stored in the entire database was 48,512,596, and the total number of characters was 3,437,337,516.

Table 1. Categories of Japanese onomatopoeia

Category	Example of Onomatopoeia	Meanings
Repetitive1	ぐいぐい /gui-gui/	strongly
Repetitive2	ふらふら /hura-hura/	woozily
Emphatic	しっかり /sikkari/	steadily
"-ri" suffixed	ばさり /basari/	readily
"-Q" suffixed	ぽたっ /potaQ/	trickly
"-N" suffixed	ぱちん /pachiN/	(with verb "shut") with the sound "snap"
Long Vowel	どーん /do:N/	reliably

* "Q" : double consonant; "N" : Japanese syllabic nasal; ":" : long vowel

We selected one search word from onomatopoeia list and from the database we acquired all the sentences that included the search word. In addition, we assigned corresponding speaker information to all acquired utterances. Finally, we calculated the total number of utterances of each speaker and the number of utterances which included onomatopoeia. This processing was conducted for all onomatopoeia search words in the database from 1992 to 2012.

2.4 Speakers' hometowns

We selected speakers who had more than 1000 utterances and investigated their hometowns or birthplaces in internet databases (Wikipedia or their official websites). As a result, the hometowns of 3,606 of 5,803 people (62.12%) who provided more than 1,000 utterances in the Diet minutes were revealed. The speakers whose hometowns were inside Japan were included in the analysis.

3 Results

As in the previous study [3], the speakers were divided into 11 categories according to their hometowns. Table 2 presents the distribution of frequent onomatopoeia users in Japan.

Table 2. Number of speakers in each home region

Home regions	Number
Hokkaido	69
Touhoku	155
Kanto	492
Koushin-etsu	99
Hokuriku	64
Toukai	155
Kinki	262
Chugoku	120
Shikoku	73
Kyusyu	189
Okinawa	20

3.1 Frequency of all onomatopoeia

Figure 1 is a heat map which expresses the frequencies of all onomatopoeia words in the list in each region using color (red: high frequency; blue: low frequency). To produce this heat map, the number of sentences with onomatopoeia in each prefecture was divided by the total number of sentences in each prefecture, and the number of sentences with onomatopoeia was divided by the total number of sentences by each speaker. Then the average numbers for each prefecture were worked out, and these numerical values were normalized with the maximum to control the bias. Finally, a heat map was rendered based

on these values.

Because the color distribution of Figure 1 is largely uniform, we concluded that there is no difference in the frequencies of onomatopoeia use among the regions in Japan. For result 2, we conducted an exploratory analysis of the data for certain categories of onomatopoeia.

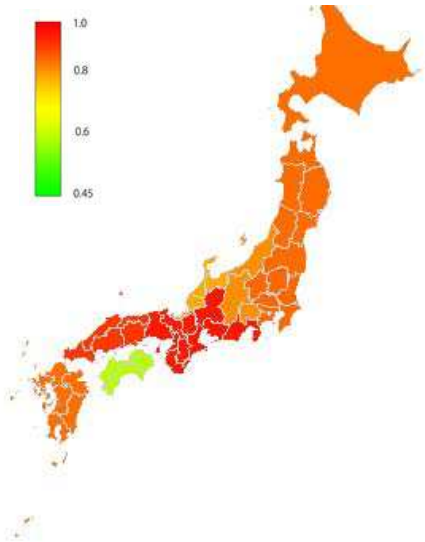


Figure 1. Frequency of onomatopoeia use (all items in list)

3.4 Frequency of emphatic onomatopoeia

Figure 2 is the heat map of “emphatic” onomatopoeia. Emphatic onomatopoeia words are in CVCCV-*ri* form (C: consonant, V: vowel): for example, *jikku-ri* (“slowly and carefully"), *hakki-ri* (“clearly"), *sikka-ri* (“well"). These onomatopoeia are the emphatic form of the basic onomatopoeia like *jikuri*, *hakiri*, *sikari*, so they are called “emphatic onomatopoeia”. For this type of onomatopoeia, the most frequent usage was found in the Kinki (shown in red).

In addition, we analyzed the frequencies of adverbs that are not onomatopoeia (for example, *jyu-bun* "sufficient" or *mata* "again") to investigate if differences in frequency among the regions is a phenomenon specific to onomatopoeia. Figure 3 is the heat map of the frequency of such adverbs. The color distribution of this map is also uniform, similar to Fig. 1. From this result, differences in the frequency of adverbs were not observed.

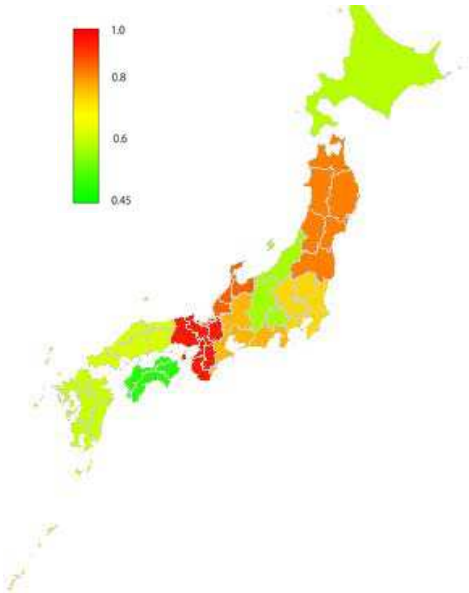


Figure 2. Frequency of emphatic onomatopoeia use

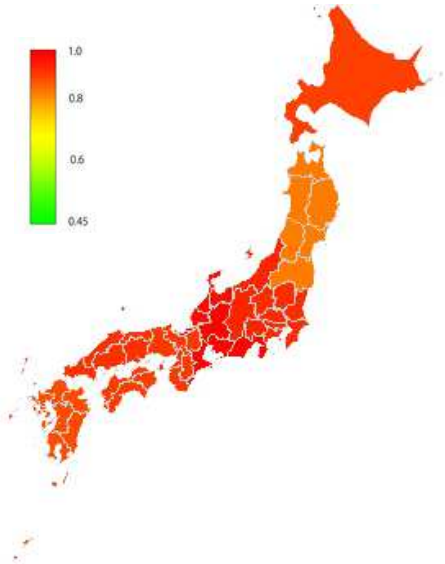


Figure 3. Frequency of non-onomatopoeic adverb use

4 Discussion

4.1 Summary

The purpose of this research was to reveal differences in the frequency of use of onomatopoeia among regions in Japan. With the records of the Diet of Japan as a massive conversation data set, we found an absence of frequency differences for all types of onomatopoeia except for the emphatic type, which was most prevalent in the

Kinki region. Moreover, this difference is a specific characteristic of onomatopoeia, because it was not observed with words that play the same role in conversation (i.e., adverbs).

The results of this research are consistent with a study which analyzed talk in local TV programs in the Kinki region [5]. This study pointed out that onomatopoeia such as "ga:Q" "ba:Q" (both also emphatic types of onomatopoeia) were frequently used. With this observation, the flok-theory that argues that "onomatopoeia is most frequently used in the Kinki area" may result from observations of a single type of onomatopoeia, rather than observations of all types. To examine this possibility, detailed research focused on the conversation data of Kinki locals is necessary.

4.2 Discussion

This study has several limitations. First, the hometowns of all the selected speakers were not clear. Using various resources we were unable to determine whether the stated hometown was just the birthplace or was the place they spent their childhood. Furthermore, even if the information from the resource was correct, it was not clear if speakers retained the dialect of their hometowns. Older speakers had a typical pattern of moving, for example, living from birth to high school age in their hometowns and then going to Tokyo at the time of university entrance. But it is now very difficult to accurately locate the place where younger people form their language identity, because people nowadays tend to wander from place from place to place, including to and from foreign countries, from childhood. To prevent these problems, dialectology usually adopts a micro-viewpoint, with detailed descriptions of a small number of speakers. The present study offered what may be called a macro-viewpoint, using large-scale data. Combining these two viewpoints may offer a way to break an impasse in both dialectology and anthropology in the future.

For generalization of this research, we can point out the importance of relationship between dialect and onomatopoeia. The research intended for the distribution of English phonesthemes in England [6] showed clear distributional patterns in certain areas. Even English, which includes less onomatopoeia (=phonesthemes) than Asian Languages, showed the distributional bias inside country. It implies that there is larger impact in Asian languages of the relationship between onomatopoeia and dialects or certain areas. It is getting more important to consider the characteristics of onomatopoeia not only in a language but also within a language to understand and use onomatopoeia efficiently. We hope that this quantitative approach will offer a new viewpoint to the dialectology and also effective use of onomatopoeia in the field of information science.

Acknowledgment

This research was funded by a Grant-in-Aid for Young Scientists (A) to Satoshi Nakamura (23680006).

References

[1] Kakehi, H., Tamori, I., and Schourup, L.: *Dictionary of Iconic Expressions in Japanese*, Berlin/New York: Mouton de Gruyter (1996)

[2] Mitsui, H., Inoue, F. The use of dialect database, Kobayashi, T. (Ed), *Technical Reports of Dialectology*, 29-89, Iwanami-Shoten, Tokyo, 2007 (in Japanese).

[3] Hirata, S., Akita, K., Komatsu, T., Nakamura, S., Fujii, H., Sawai, D. A dialectal comparison of the image and use of onomatopoeia, *The 26th Annual Conference of the Japanese Society for Artificial Intelligence*, 2012a (in Japanese).

[4] Hirata, S., The frequency of use of onomatopoeia in daily conversations, *The 76th Annual Conference of the Japanese Psychological Association*, 2012b (in Japanese).

[5] Tahara, H., Pyatto chigitte syatto watasu -Kansaiben no onomatope- (I tear it off with “PyaQ” and hand it with “SyaQ” - Onomatopoeia in Kansai Dialect). *Gekkan Gengo*, 30, 24-25, 2001 (in Japanese).

[6] Wright, D., Scrunch, growze, or chobble?: investigating regional variation in sound symbolism in the Survey of English Dialects. *Leeds Working Papers in Linguistics and Phonetics* 17, 2012.

The Second Asian Conference on Information Systems, ACIS 2013, aims to bring together many scholars, researchers and managers of various areas and industries for intellectual exchanges, research cooperation and professional development. This conference features two tracks inviting presentations of theoretical research findings and case reports. Moreover, this program also offers excellent networking opportunities to participants, with a wonderful taste of local culture. High quality research papers and case studies are invited to be submitted electronically through the conference's website. The conference areas of interest include, but are not limited to, the following:

- Artificial Intelligence
- Data/Text Mining
- Communication Technology
- Image Processing
- Pattern Recognition
- Information/Cyber Security
- Information Retrieval
- Information Extraction
- Intelligent Systems and Agents
- Natural Language Processing
- Ubiquitous Computing
- RFID Technology
- E-health, E-education
- E-government System
- Enterprise Information Systems
- Decision Support Systems
- ICT For Development (ICT4D/ICTD)
- Information Security
- Encryption Technologies
- Knowledge-based System
- Software Engineering
- Learning-Supported System
- Web-based Intelligent Applications
- Innovative Business Models
- Service Design and Engineering
- Service Science and Management
- Service Innovation and Marketing
- E-Services and Service Computing
- Human Resource Service
- Knowledge complexity and metrics
- Knowledge management
- Knowledge representation/reasoning
- Knowledge verification/validation
- Logistics Management and Innovation

Title: Proceedings of of the Second Asian Conference on Information Systems(ACIS 2013)
ISBN: 978-616-91886-0-5

Contact: Sirindhorn International Institute of Technology, Thammasat University
131 Moo 5 Tiwanont Road, Bangkadi, Muang Pathumthani 12000 Thailand
Tel: +66-2-501-3505(-20)
Fax: +66-2-501-3524
e-mail: thanaruk@siit.tu.ac.th
URL: saki.siit.tu.ac.th/acis2013