

Tai-hoon Kim Hojjat Adeli
Jianhua Ma Wai-chi Fang
Byeong-Ho Kang Byungjoo Park
Frode Eika Sandnes Kun Chang Lee (Eds.)

Communications in Computer and Information Science

264

U- and E-Service, Science and Technology

International Conference, UNESST 2011
Held as Part of the Future Generation
Information Technology Conference, FGIT 2011
in Conjunction with GDC 2011
Jeju Island, Korea, December 2011, Proceedings



Springer

Tai-hoon Kim Hojjat Adeli Jianhua Ma
Wai-chi Fang Byeong-Ho Kang
Byungjoo Park Frode Eika Sandnes
Kun Chang Lee (Eds.)

U- and E-Service, Science and Technology

International Conference, UNESST 2011
Held as Part of the Future Generation
Information Technology Conference, FGIT 2011
in Conjunction with GDC 2011
Jeju Island, Korea, December 8-10, 2011
Proceedings

Volume Editors

Tai-hoon Kim

Hannam University, Daejeon, Korea

E-mail: taihoonn@hannam.ac.kr

Hojjat Adeli

The Ohio State University, Columbus, OH, USA

E-mail: adeli.1@osu.edu

Jianhua Ma

Hosei University, Tokyo, Japan

E-mail: jianhua@hosei.ac.jp

Wai-chi Fang

National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

E-mail: wfang@mail.nctu.edu.tw

Byeong-Ho Kang

University of Tasmania, Hobart, TAS, Australia

E-mail: byeong.kang@utas.edu.au

Byungjoo Park

Hannam University, Daejeon, Korea

E-mail: bjpark@hnu.kr

Frode Eika Sandnes

Oslo University College, Norway

E-mail: frode-eika.sandnes@hioa.no

Kun Chang Lee

Sungkyunkwan University, Seoul, Korea

E-mail: kunchanglee@gmail.com

ISSN 1865-0929

e-ISSN 1865-0937

ISBN 978-3-642-27209-7

e-ISBN 978-3-642-27210-3

DOI 10.1007/978-3-642-27210-3

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: Applied for

CR Subject Classification (1998): I.2, H.3, H.4, C.2, F.1, H.5

© Springer-Verlag Berlin Heidelberg 2011

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

u- and e-service, science and technology are areas that attract many professionals from academia and industry for research and development. The goal of the UNESST conference is to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions relating to the multifaceted aspects of u- and e-service, science and technology.

We would like to express our gratitude to all of the authors of submitted papers and to all attendees for their contributions and participation.

We acknowledge the great effort of all the Chairs and the members of Advisory Boards and Program Committees of the above-listed event. Special thanks go to SERSC (Science and Engineering Research Support Society) for supporting this conference.

We are grateful in particular to the speakers who kindly accepted our invitation and, in this way, helped to meet the objectives of the conference.

December 2011

Chairs of UNESST 2011

Preface

We would like to welcome you to the proceedings of the 2011 International Conference on u- and e-Service, Science and Technology (UNESST 2011)—the partnering event of the Third International Mega-Conference on Future-Generation Information Technology (FGIT 2011) held during December 8–10, 2011, at Jeju Grand Hotel, Jeju Island, Korea.

UNESST 2011 focused on various aspects of advances in u- and e-service, science and technology. It provided a chance for academic and industry professionals to discuss recent progress in the related areas. We expect that the conference and its publications will be a trigger for further related research and technology improvements in this important subject.

We would like to acknowledge the great effort of the UNESST 2011 Chairs, Committees, International Advisory Board, Special Session Organizers, as well as all the organizations and individuals who supported the idea of publishing this volume of proceedings, including the SERSC and Springer.

We are grateful to the following keynote, plenary and tutorial speakers who kindly accepted our invitation: Hsiao-Hwa Chen (National Cheng Kung University, Taiwan), Hamid R. Arabnia (University of Georgia, USA), Sabah Mohammed (Lakehead University, Canada), Ruay-Shiung Chang (National Dong Hwa University, Taiwan), Lei Li (Hosei University, Japan), Tadashi Dohi (Hiroshima University, Japan), Carlos Ramos (Polytechnic of Porto, Portugal), Marcin Szczuka (The University of Warsaw, Poland), Gerald Schaefer (Loughborough University, UK), Jinan Fiaidhi (Lakehead University, Canada) and Peter L. Stanchev (Kettering University, USA), Shusaku Tsumoto (Shimane University, Japan), Jemal H. Abawajy (Deakin University, Australia).

We would like to express our gratitude to all of the authors and reviewers of submitted papers and to all attendees, for their contributions and participation, and for believing in the need to continue this undertaking in the future.

This work was supported by the Korean Federation of Science and Technology Societies Grant funded by the Korean Government.

December 2011

Tai-hoon Kim
Hojjat Adeli
Jianhua Ma
Wai Chi Fang
Byeong-Ho Kang
Byungjoo Park
Frode Eika Sandnes
Kun Chang Lee

Mei-Ling Shyu
Nguyen Manh Tho
Rami Yared
Raymond Choo
Regis Cabral
Sajid Hussain

Seng W. Loke
SeongHan Shin
Sheng Zhong
Simone Fischer-Hübner
Stefano Ferretti
Tingting Chen

Weijia Jia
Yao-Chung Chang
Young Jin Nam
Zhaohao Sun Hebei
Zhenjiang Miao
Zhuowei Li

Special Sessions Organizers

Hee Kyung Kong
Kun Chang Lee
Hanmin Jung
Yung-Cheol Byun

Table of Contents

A Self Evolutionary Rule-Base	1
<i>A.M. Khattak, Z. Pervez, W.A. Khan, S.Y. Lee, and Y.K. Lee</i>	
Hybrid Storage Architecture for Networked Digital Signage Systems	10
<i>Youngjin Nam, Soon hwan Jeong, and Young kyun Park</i>	
Adoption of Technology Applications in Healthcare: The Influence of Attitude toward Knowledge Sharing on Technology Acceptance in a Hospital	17
<i>Pouyan Esmaeilzadeh, Murali Sambasivan, Naresh Kumar, and Hossein Nezakhati</i>	
Generic Evolutionary Framework for Simulating Business Processes	31
<i>Veng-Ian Chan and Yain-Whar Si</i>	
Test-Activity Flow Control for SOA Execution Test	42
<i>Youngkon Lee</i>	
Business-Rating Based SLA Model for SOA Quality	50
<i>Youngkon Lee</i>	
Semi-automatic Test Assertion Transformation Scheme for SOA	58
<i>Youngkon Lee</i>	
Idea Sketch of Application Service of e-Business Card and NFC Based on Mobile	66
<i>Sun Kuh Noh, Byung Rae Cha, Dong You Choi, Jung-Jin Park, Guen Sik Kim, and Yoo Kang Ji</i>	
Framework of Design Patent Map Systems Based on Dissimilarity from Populace Survey	75
<i>Chao-Chun Chen, Rain Chen, and Ting-Ting Dai</i>	
Service Quality Assessment in Provision of Library Services	83
<i>Masitah Ahmad, Jemal Abawajy, and Tai-hoon Kim</i>	
An Adaptive IEEE 802.15.4 Data Transmission Scheme for Smart Grid Applications	93
<i>Zhongzhi Che, Jianhua He, Zuoyin Tang, Yi Zhou, and Kai Chen</i>	
Modeling Impact of Both Frame Collisions and Frame Corruptions on IEEE 802.15.4 Channel Access for Smart Grid Applications	100
<i>Zhongzhi Che, Jianhua He, Yi Zhou, Zuoyin Tang, and Chao Ma</i>	

Training Schedule Management Application Using Various Constraints	106
<i>Orawan Mantrawong and Chantana Chantrapornchai</i>	
Data Mining and Web-Based Children Shoe Suggestion System	114
<i>P. Sudta, K. Kanchan, and Chantana Chantrapornchai</i>	
Proper Parallel Numerical Integration Method for Magnetic Sounding Equations on Multilayered Earth.....	122
<i>Banpot Dolwithayakul, Chantana Chantrapornchai, and Suabsakul Yooyeunyong</i>	
QoS Guarantee with Adaptive Transmit Power and Message Rate Control for DSRC Vehicle Network Based Road Safety Applications ...	130
<i>Wenyang Guan, Jianhua He, Lin Bai, Zuoyin Tang, and Yi Zhou</i>	
The Influence of Wine Expert's On-Line Ratings on Retail Price of On-Premise Market in Korea	136
<i>YoonJung Nam and Youngsik Kwak</i>	
Shipment Timing Support System Modeling for Stored Apples in Korea	144
<i>Youngsik Kwak, Seokil Song, and Yoonsik Kwak</i>	
Heuristic for Paper Roll Selection for Corrugation Process Simulation...	153
<i>Chantana Chantrapornchai and T. Sathapanawat</i>	
A Spatial Morphology of the English Verb Structure of Simple Sentences Aids Appropriate Cognitive Structures Develop and Offers EFL/ESL Potential as a Mobile App	160
<i>Robert C. Meurant</i>	
An Empirical Analysis of the Effects of IT Professionals' Emotional Dissonance on Creativity Revelation Processes and Individual Creativity	168
<i>Seong Wook Chae, Young Wook Seo, and Kun Chang Lee</i>	
The Impact of Creative Self-efficacy, IT Support, and Knowledge on Individual Creativity through Absorptive Capacity	177
<i>Young Wook Seo, Seong Wook Chae, and Kun Chang Lee</i>	
An Empirical Analysis of Leadership Styles and Their Impact on Creativity: Emphasis on Korean ICT Companies	187
<i>Nam Yong Jo, Dae Sung Lee, and Kun Chang Lee</i>	
A Physiological Approach to Creativity under Stress and Non-stress Conditions	197
<i>Dae Sung Lee, Nam Yong Jo, and Kun Chang Lee</i>	

The Structure of Individual Creativity Revelation Processes with Task Characteristics and Social Characteristics: An Empirical Analysis Based on IT-Services Participants	207
<i>Do Young Choi, Min Hee Hahn, and Kun Chang Lee</i>	
Effects of Social and Emotional Intelligence on the Creative Process and Individual Creativity	217
<i>Min Hee Hahn, Do Young Choi, and Kun Chang Lee</i>	
An Management of Digital Contents on Science and Technology.....	227
<i>Won-Goo Lee, Min-Ho Lee, Kwang-Young Kim, Sung-Ho Shin, Hwa-Mook Yoon, and Won-Kyung Sung</i>	
A Categorization and Solution of Metadata Conversion Problem on Research Paper	230
<i>Min-Ho Lee, Hwa-Mook Yoon, Won-Kyung Sung, and Jae-Cheol Ryou</i>	
Multi-words Terminology Recognition Using Web Search	233
<i>Sa-Kwang Song, Yun-Soo Choi, Hong-Woo Chun, Chang-Hoo Jeong, Sung-Pil Choi, and Won-Kyung Sung</i>	
Disambiguating Author Names Using Automatic Relevance Feedback ...	239
<i>In-Su Kang and Seung-Hoon Na</i>	
OntoURIResolver: URI Resolution and Recommendation Service Using LOD	245
<i>Taehong Kim, Pyung Kim, Seungwoo Lee, Hanmin Jung, and Won-Kyung Sung</i>	
A Comparison of Classifiers for Detecting Hedges	251
<i>Sin-Jae Kang, In-Su Kang, and Seung-Hoon Na</i>	
Application of Trend Detection of Technical Terms to Technology Opportunity Discovery	258
<i>Mi-Nyeong Hwang, Min-Hee Cho, Myung-Gwon Hwang, Mikyoung Lee, and Do-Heon Jeong</i>	
Decision-Making Support Service Based on Technology Opportunity Discovery Model	263
<i>Mikyoung Lee, Seungwoo Lee, Jinhjung Kim, Dongmin Seo, Pyung Kim, Hanmin Jung, Jinhee Lee, Taehong Kim, Hee Kwan Koo, and Won-Kyung Sung</i>	
Relation Extraction Based on Composite Kernel Combining Pattern Similarity of Predicate-Argument Structure	269
<i>Hong-Woo Chun, Chang-Hoo Jeong, Sa-Kwang Song, Yun-Soo Choi, Sung-Pil Choi, and Won-Kyung Sung</i>	

English-to-Korean Cross-Lingual Link Detection for Wikipedia.....	274
<i>Ralph Marigomen and In-Su Kang</i>	
Efficient Finding Relationship between Individuals in a Mass Ontology Database	281
<i>Dongmin Seo, Hee Kwan Koo, Seungwoo Lee, Pyung Kim, Hanmin Jung, and Won-Kyung Sung</i>	
Generating Knowledge Map for Acronym-Expansion Recognition	287
<i>Do-Heon Jeong, Myung-Gwon Hwang, and Won-Kyung Sung</i>	
Inference Verification for Enhancing Confidence in Semantic Information Service	294
<i>Seungwoo Lee, Mikyoung Lee, Hanmin Jung, and Won-Kyung Sung</i>	
Design of TOD Model for Information Analysis and Future Prediction	301
<i>Jinhyung Kim, Seungwoo Lee, Jinhee Lee, Mikyoung Lee, and Hanmin Jung</i>	
A the Effects of SCM Activities on Manufacturing Capabilities and Manufacturing Performances	306
<i>Su-Hyun Jung, Sung-Hong Kim, and Sung-Lyong Kang</i>	
A Digital Convergence Maturity Model: The Relative Importance of Factors	316
<i>Seung-Jun Yeon and Sung-Hyun Hwang</i>	
Examining Consumer Preferences for Mobile Devices to Read Digital Content for the Diffusion of Ubiquitous Learning in Higher Education.....	324
<i>Hyeongjik Lee, Won Bin Lee, and Soo Cheon Kweon</i>	
A Service Quality Model for the Public Information Service	332
<i>Hyun-Woong Jo and Sang-Wook Kim</i>	
A Study on the Smartwork Policy for the Productivity in Ageing Advanced Economies	341
<i>Hee-kyung Kong and Kihoon Sung</i>	
Erratum	
Idea Sketch of Application Service of e-Business Card and NFC Based on Mobile	E1
<i>Sun Kuh Noh, Byung Rae Cha, Dong You Choi, Jung-Jin Park, Guen Sik Kim, and Yoo Kang Ji</i>	
Author Index	347

A Self Evolutionary Rule-Base

A.M. Khattak, Z. Pervez, W.A. Khan, S.Y. Lee, and Y.K. Lee

Department of Computer Engineering, Kyung Hee University, Korea
{asad.masood, zeeshan, wajahat.alikhan, sylee}@oslab.khu.ac.kr,
yklee@khu.ac.kr

Abstract. The rapid growth in domain knowledge is the main reason for the evolution of knowledgebase's maintaining the domain knowledge. Rule-based Decision Support Systems (DSS) are the most effected systems with the growing knowledge. The experts need to continuously update the rule-base for the new knowledge. This manual and periodic updates in rule-base are time consuming and less useful. In this paper we propose a Self Evolutionary Rule-base algorithm for rule-bases of DSS to decrease the burden from experts and also provide updated knowledge on time. To achieve this objective, we develop a generic structure for rules storage that not only provide efficient manipulation of rules but a generic structure for storage of rules regardless of rules nature/format. The detail working of proposed Rule-base system for rules storage and manipulation is provided in this paper. For the proof of concept, we have implemented the Self Evolutionary Rule-base algorithm in Socially Interactive Clinical Decision Support System (SI-CDSS). The focus is on diabetes disease patients and the overall SI-CDSS is deployed in Microsoft Azure environment. In its implementation, Rough Set generated rules are used and the algorithm is executed on Rough Set generated rules.

Keywords: Rules, Rule-base, Decision Support System, Algorithms, Performance, Verification.

1 Introduction

Rapid growth of information is the main reason for the increase in complexity of the collected data in a specified field of study. These complexities also introduce various types of uncertainties in the data collected specifically relating to problems in healthcare applications and services [2 and 3]. To extract the useful information from the uncertain data, different researchers from mathematics, computer science, and medical related areas have worked on number of theories [10] that supports in building expert systems. However, these theories do have limitations. These limitations restrict the theories and its complaint systems to limited domains where they can perform better.

Rules are the important paradigm for representing expert knowledge. In rule-based expert systems, the rule-base contains the domain knowledge coded in the form of rules. In healthcare domain, theories that support rule-base models are preferred over those of machine learning based theories [2]. The fact behind it is that medical doctors

can interact with systems using rule-based approaches. They can compile new rules while in machine learning based approaches, medical doctors are not aware of the internal working of the systems. In addition, most of the domain specific expert systems are also based on rule-base (knowledgebase) [11]. Rule-based systems store, manipulate, and interpret information in a useful way. Rule-based systems are frequently used for diagnosis, recommendations, and symptoms clustering [8].

The rule-base systems use rules for variety of different purposes such as conflict resolution, decision making, and recommendation. However, the main issue is that these rules are maintained in text files or in XML files and are static in nature; on the other hand, the domain these rules are defined for is not static. So there is a need for a system that can help in rule evolution with the evolution in the field of knowledge. Rough Set based technique is used to generate new rules from the data collected [3]. The new rules need to update the existing set of rules or in other words, making the rule-base to evolve so that the rule-base can accommodate the new or changed rules. Another issue is a common structure for the storage of rules. When a system uses different rule-based approaches (rules in different format) then for rules storage they also need different structures.

In this research, we focus on a generic storage structure that is used for storage of rules (from healthcare domain) of different representational format. On top of the proposed storage structure, we propose a Self Evolutionary Rule-base algorithm for rules with changes reflected from the evolved domain knowledge to accommodate the new and changed knowledge in Rule-base. A detail discussion on storage structure for rules and the implementation of Self Evolutionary Rule-base algorithm is provided in this research. The proposed algorithm for Rule-base is implemented with our underdevelopment Socially Interactive Clinical Decision Support System (SI-CDSS) [2] shown below.

The remaining paper is arranged as follows: Related Work section presents details on Rule-base and Rule-based systems. Section Proposed Scheme provides detail description of proposed storage structure and evolutionary algorithm. Implementation and Results section presents the preliminary implementation details. Finally we conclude our discussion in Conclusions section.

2 Related Work

Expert systems are mostly rule-based systems. They use human expert knowledge (human intelligence) to solve real-world problems. The expert knowledge is often represented in the form of rules. Rule-based systems in healthcare domain are preferred over machine learning based systems due to the fact that rules are easily understandable to doctors. The rules in healthcare system represent medical doctors and domain knowledge. In addition, medical doctors feel comfortable while interacting with rule-based systems. The problem with these systems is that the rules are static and needs expert intervention for updates [3]. A medical guidelines based (i.e., rule-based) clinical decision support system proposed in [4] mainly focus on metabolism synthesis. The rule-base is constructed from the predefined logic (i.e.,

medical logic) and then used in inference engine for decision making. It is composed of four different components namely; data, model, inference, human computer interaction. The rule-base is static and does not consider evolution for the updates in the field of knowledge (domain knowledge). So, all the updates are carried out by experts and doctors manually.

In [7], the authors focused on real-time activities performed by patients and then use these activities and domain knowledge for decision making. The system used description logic rules after match making process to make appropriate decisions based on situation analysis. The rules used are stored in text file and are periodically updated by experts based on user and system needs. In [5], the authors proposed inference mechanism with Electronic Health Record (EHR) for existing hospital information system. The rule engine and the rule-base for containing the clinical guideline are integrated together. With static rules, their main drawback is strict modeling of information according to its input types for different components.

A healthcare service using conventional clinical decision support system and ontology to manage user healthcare data has been proposed in [6]. The main functionality of the system was to generate, deploy, and manage patient information. The information is inferred using rule-base and ontology. The results are then propagated to dependent components. In [9], the authors presented a tool for multi-objective job scheduling problems. An interactive multi objective genetic algorithm for decisions has been proposed. It's decision function is defined as a measure of truth of a linguistically quantified statements (rules). The tool also provides support for what-if analyses.

Research work and actual deployment of rule-based decision system is visible in different areas including healthcare such as in development of clinical decision support systems. These systems are very strict with the structure for rule representation and one rule-base cannot support two or more inference engines with different formats of rules. Another issue is that the rule-base of these systems is static and is not updated with the new knowledge. Our proposal is a generic structure for rules storage regardless of the rules nature and keeps the rule-base dynamic with evolving domain knowledge.

3 Proposed Scheme

Clinical Decision Support System (CDSS) is widely used now days in every country for better, timely, and low cost healthcare services. Most of these systems are based on the concept of using rules for decision making and because of this these systems are also termed as expert systems. This section provides details on proposed scheme i.e., Self Evolutionary Rule-base for SI-CDSS [2]. SI-CDSS continuously generate new rules based on patient's experience and expert interventions. To cope with the growing knowledge, our proposed scheme keeps the rule-base updated. To support the proposed scheme, a generic structure is also build for rules storage regardless of rules structure/format. Based on the generic structure and evolutionary algorithm, this section is divided into two subsections (see Figure 4 for overall system architecture).

3.1 Rules-Base Storage Structure

Information systems as well as decision support systems are mainly based on reliable data and facts. Facts (conditions) represent our knowledge about the situation/problem. Rules are used to represent relationships among the Facts. Based on these rules, Inference Engine makes the inference for situation analysis and decision making. The rules are mainly used to get end results based on the given facts. For example, Figure 1 shows the consequent effects of combination of facts (A....X) that builds confidence for a rule in a given situation and results in consequent effects (actions).

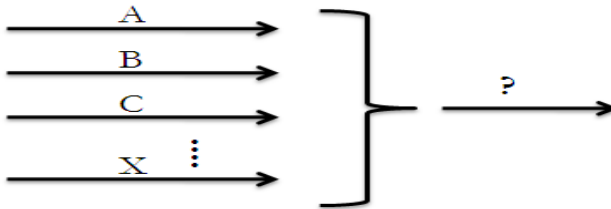


Fig. 1. Shows abstract structure of rule. Set of conditions (facts), confidence building, and resultant action.

In most systems these rules are stored in text files that are hard to maintain. In case of systems using two or more inference engines such as; SI-CDSS [2], the systems will maintain two different rule-bases as the rules format is different (see Figure 2) and cannot be stored in a rule-base fix for one particular format. As it is obvious that rules have two main components i.e., condition and action that makes it easy to model it. With variety of different rules representational format, using different logical, mathematical, and relational operators makes it a bit tough to model them in a single storage structure. To handle this, we develop a schema and in that we model rule as an entity separate from its condition and action parts. The rule entity contains the information about its associated conditions and actions information by distributing its key as a foreign key. On rule entity, it also contains information about the rule format and the type of disease it is focusing on. All the conditions that can be part of a rule

if BMI<=3, Lethargic=1, Fats=1, then Diabetes type=1 if Exercise>=2, Fats=0, then Additional Medication=0
Patient(p) □ hasSymptoms(cool skin) □ hasSymptoms(dam skin) □ hasBP(160/120) →generateRecomendation(drink juice) □ generateRecomendation(take rest)
UrineProblem,Weight,WeightLoss,Age,Fatigue,Pain→Class Yes,73,*,48,No,* → 5

Fig. 2. Shows the example of procedural (if-then) rules, description logic rules, and rules used by Rough Set

are modeled separately along with the operators. This also solves another problem for storing and handling unknown number of conditions of a rule while in conventional systems, the numbers of conditions are fixed. The same way, the actions can be many for a given rule so we also model it in a separate entity that makes a many-to-one relationship with the rules entity. For more detailed view of the storage structure please see Figure 3. The class entity is for extra recommendations from medical doctors while the user entity is for tracking user and user access privileges.

3.2 Self Evolutionary Rule-Base

Rule-base is one of the main components of SI-CDSS that store and manipulate decision rules based on user's request. Inference Engine needs to take a decision and for decision making rule-base is contacted and corresponding rules are extracted for decision making. These rules are mainly composed based on experts (doctors) knowledge [1], however, in SI-CDSS, patient's experience is also incorporated that further help in recommendations and decision making. As updates in the domain knowledge and patients experience sharing is very frequent, and is the main reason for frequent rules generation by Rough Set based inference engine. At the same time these rules need to be updated in the Rule-base to provide up to date services and recommendations.

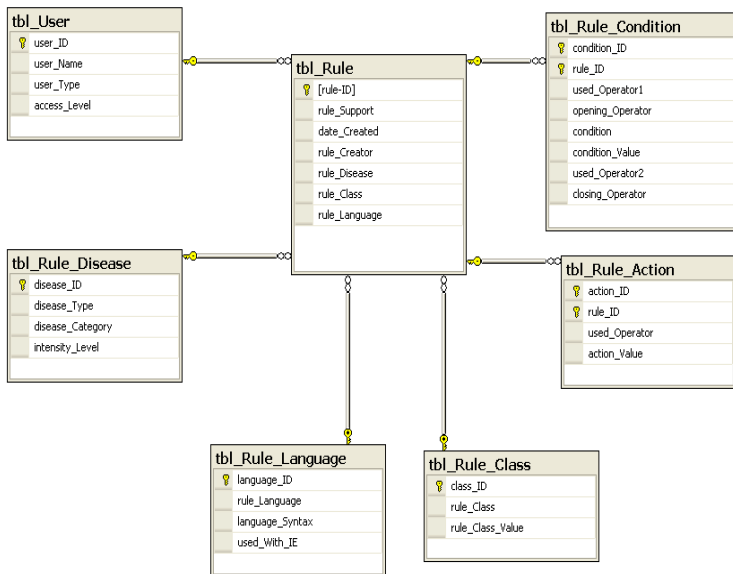


Fig. 3. Shows the generic storage structure for rules storage in rule-base

To achieve better healthcare services, recommendations, and decision makings; more sophisticated and exhaustive list of rules that can serve better are required. With the passage of time, the advancement in expert knowledge and user experience may

introduce new rules as well as changes in the existing rules. These all need to be accommodated appropriately in the Rule-base. Change in rules based on patient's experience is very sensitive. For this reason, the generated rules will only be stored if they are verified by the experts and allowed to be stored in the repository. We use the notion of social as the system is more interactive and also use patient's experience for recommendations and decision makings. These rules are used by inference engine for decision making or analysis, so its main interaction point is the inference engine. Updates in rules and Rule-base will all happen through inference engine. The complete process of rules parsing, addition, and updates is given in the system overall architecture shown in Figure 4.

Wrapper component is responsible for selecting the appropriate rules from the rule-base regarding its format and the calling inference engine. Parser handles all the rules retrieval and submission requests. Rule Engineering and Verification is responsible for providing facility to experts for creating new rules and also verification of existing rules in the Rule DB of Rule-base. Rule Updates component is actively listing for changes in existing rules. Once change is found then it update the corresponding rule and propagate it to Rule DB. It is important to note that not all the rules are updated based on simple change in rules but in fact, a rule is updated if there is a change in the consequent (action) of the rule. The update can be a change of existing consequent or addition of another consequent and in our proposed structure we have provision for any number of additional consequent. The detail working of proposed Self-Evolutionary Rule-base system is given in Algorithm 1.

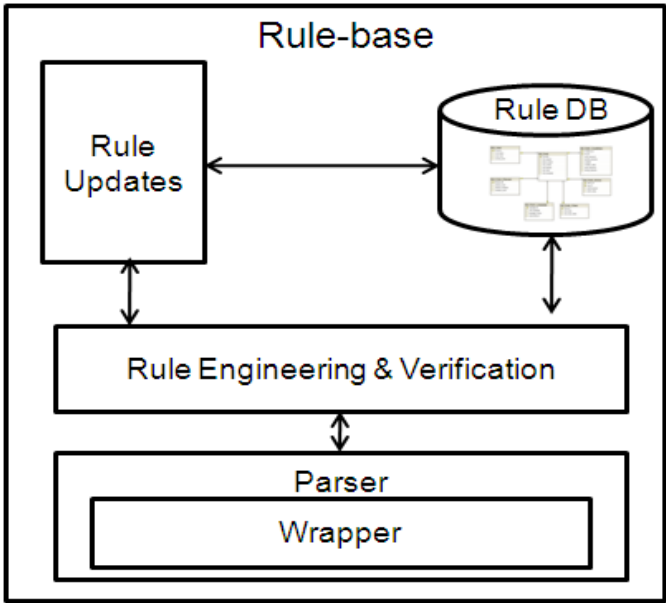


Fig. 4. Overall architecture diagram of proposed Self Evolutionary Rule-base

Algorithm EvolutionaryRule-base (): Self Evolutionary algorithm to evolve Rule-base for accommodating the updated knowledge.

Input: Rules form Rule DB and newly generated rules from inference engine.

Input: User entered rules for new knowledge.

Output: Set of new and updated rules storage.

```

1. /* Check for type of inference engine and then activate appropriate wrapper.*/
2.  Wrapper.initiate(IE.RulesType)
3. /* Fetch the rules generated by the inference engine or expert entered rules.*/
4.  Rules ← IE.Rules
5.  Rules + ← EXPERT.Rules
6. /* Fetch rules from RuleDB to be updated.*/
7.  Rulesdb ← RuleDB.Rules
8. /* Check if rules from inference engine are new then add in RuleDB.*/
9.  IFNOT(Rules= ∃Rulesdb)then
10. Rulesdb ← {x | <RulesΔ, x >New}
11. Endif
12. /* Check for rules updates, update rule, and store in RuleDB*/
13. IF(Rules= ∃RulesDB) ∩ (∃Δ ∩ Δ.Rules.Change)then
14. Rulesdb.Rule ← {x | <RulesΔ, x > Change}
15. Endif
16. /* Update the original RuleDB for the new and updated rules.*/
17. Execute.update(RulesDB, Rulesdb)
18. End

```

Algorithm 1. A Self Evolutionary Rule-base algorithm for dynamic updates in Rule DB for evolving domain knowledge

4 Implementation and Results

As mentioned above, the Rule-base system is a subcomponent of overall SI-CDSS [2], which has been developed and deployed on Microsoft Azure environment for its prototype demonstration. In SI-CDSS deployment, Rule-base is also implemented and deployed for Rough Set based inference engine with focus on diabetes disease. In this prototype demonstration, the Rule DB was developed and maintained in Microsoft SQL Azure. Figure 5-a shows the user interface for patients to share their experience of using the medicine. This information is then used by Rough Set for rules generation and after mining the experience, inference engine generates recommendations for the patients as shown in Figure 5-b. The rules used by Rough Set are also in different format than the other rules see Figure 2. All the attributes used in conditions and actions are listed as columns, are of fix number. The values for the conditions are separated using comma that represent the AND operation. The rules contain different values for different combinations of symptoms. The numbers of symptoms are 33 in every single rule. With system learning new rules, we test our system for its claims using proposed structure with database manipulation against text files using string manipulations. We introduce a bunch of random 20 rules (i.e., new rules, existing rules, and some updated rules) and test both techniques performance. All these testing and deployment experiments are carried out on local machine with 3 GB memory and 2.67 GHz of quad core processor.

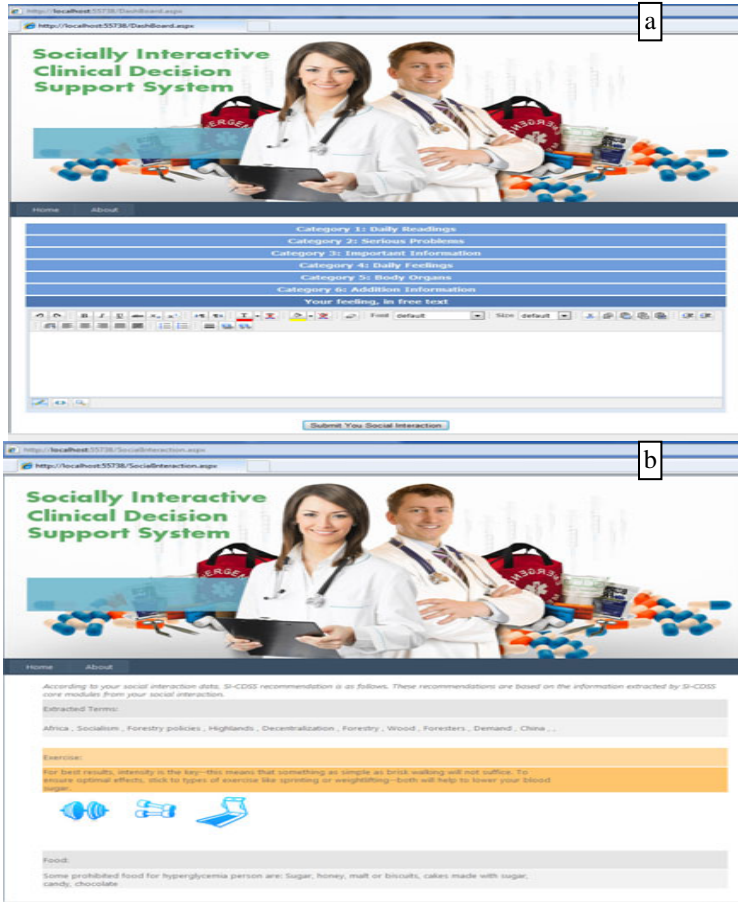


Fig. 5. (a) shows the interface for patient's experience entry while (b) shows the recommendations generated by SI-CDSS based on the rules from Rule-base

With the novelty of introducing Self Evolutionary Rule-base, in addition, our proposed structure for rule storage also proved better performance than the traditional text file based rules storage. The Self Evolutionary Rule-base not only helps in timely updates in the Rule-base but the rules generated also become more compact. The Self Evolutionary Rule-base process is completely automatic; however, for generated rules verification, the system also have provision for experts (doctors) to verify rules and eliminate rules that are not compliant with standard knowledge of the domain.

5 Conclusion

Rule-based systems are mostly used in healthcare domain for developing CDSS. With the passage of time, the rules containing expert and domain knowledge needs to be

updated to accommodate new discoveries. To achieve this, a Self Evolutionary Rule-base system is proposed with a generic structure to store different types of rules. The system is implemented as subcomponent in currently under development system SI-CDSS in Microsoft Azure environment for diabetes patients. The overall working and performance of proposed scheme proves better in comparison against the text based rules management systems performance. In future, our focus is to work for full capacity of the proposed Rule-base i.e., to work with several different inference engines with diverse nature of rules in the Rule-base at a time and still achieve its claimed effectiveness.

Acknowledgments. This work was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agent)" (NIPA-2011-(C1090-1121-0003).

References

1. Berner, E.S., Ed, D.: Clinical Decision Support Systems: State of the Art. Department of Health Services Administration. University of Alabama at Birmingham AHRQ Publication No. 09-0069-EF (June 2009)
2. Fatima, I., Fahim, M., Guan, D., Lee, Y.K., Lee, S.Y.: Socially Interactive CDSS for u-Life Care. In: The 5th ACM International Conference on Ubiquitous Information Management and Communication (ACM ICUIMC 2011), Seoul, Korea, February 21-23 (2011)
3. Feng, F., Liu, X.Y., Leoreanu-Fotea, V., Jun, Y.B.: Soft sets and soft rough sets. *Inform. Sci.* (2010), doi:10.1016/j.ins.2010.11.004
4. Jegelevicius, D., Krisciukaitis, A., Lukosevicius, A., Marozas, V., Paunksnis, A., Barzdiukas, V., Patasius, M., Buteikiene, D., Vainoras, A., Gargasas, L.: Network Based Clinical Decision Support System. In: 9th International Conference on Information Technology and Applications in Biomedicine (ITAB), Larnaca, Cyprus (2009)
5. Kim, J.A., Cho, I., Kim, Y.: Clinical Decision Support System Architecture in Korea. In: International Conference on Convergence and Hybrid Information Technology, Washington DC, USA (2008)
6. Ko, E.J., Lee, H.J., Lee, J.W.: Ontology and CDSS based Intelligent Health Data Management in HealthCare Server. *World Academy of Science, Engineering and Technology* (2007)
7. Khattak, A.M., Pervez, Z., Ho, K.K., Lee, S.Y., Lee, Y.K.: Intelligent Manipulation of Human Activities using Cloud Computing for u-Life Care. In: The 10th Annual International Symposium on Applications and the Internet (SAINT 2010), Seoul, Korea (2010)
8. Mauno, V., Crina, S.: Medical Expert Systems. *Current Bioinformatics* 3, 56–65(10) (2008), doi:10.2174/157489308783329869
9. Petrovic, D., Duenas, A., Petrovic, S.: Decision support tool for multi-objective job shop scheduling problems with linguistically quantified decision functions. *Decision Support Systems* 43(4), 1527–1538 (2007)
10. Radzikowska, A., Kerre, E.E.: A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems* 126(2), 137–155 (2002)
11. Tsumoto, S.: Automated induction of medical expert system rules from clinical databases based on rough set theory. *Information Sciences* 112, 67–84 (1998)

Hybrid Storage Architecture for Networked Digital Signage Systems

Youngjin Nam, Soon hwan Jeong, and Young kyun Park

School of Computer & Information Technology, Daegu University, Korea
{youngjin.nam, kaoops12, youngkyunp}@gmail.com

Abstract. Digital signage has recently been highlighted as one of the next-generation digital media. It plays a given series of various quality multimedia contents after downloading them into its local disk of limited size from large-sized storage over WAN (like FTP server). In other words, the beginning of its play time can be delayed to some extent if the contents download time becomes longer. In this paper, we propose hybrid storage architecture for networked digital signage to speed up the contents download time. The proposed architecture not only includes an iSCSI storage layer between its local disk and WAN-based large-sized storage, but also provides a novel content caching algorithm. Extensive performance experiments on a simulator and its prototype with various workloads show that the proposed architecture outperforms other possible architectures in terms of the contents download time by three times.

Keywords: digital signage, storage, multimedia contents, caching, iSCSI.

1 Background and Motivation

Digital signage is referred to as a smart electronic display system to deliver various information and advertisement to anonymous users (consumers) using well-understandable multimedia contents (with overlaid texts) through 2/3D LCD and LED display devices [1][2]. In general, digital signage is connected to private or public network. Consequently, any cooperation between different digital signage systems and contents/system controls in remote could be possible. As shown in Figure 1, a typical digital signage consists of a management server, a set-top-box (STB) and a decent display device. The management server schedules a playback of a list of multimedia contents at the STB by determining a list of multimedia contents (briefly playlist), a playing order, and a starting time, and the like [3]. It also transmits the playlist information to the STB through the network. The STB is mainly in charge of playing a given playlist. Before playing them, the STB needs to have all the multimedia contents in its local disk, typically taking a considerable amount of time to download all of them from the WAN-based storage (briefly WAN storage) like FTP servers. Note that the local disk cannot hold the entire set of multimedia contents stored in the WAN storage because of its size limitation (maximally 200GB). However, the local disk can provide enough I/O bandwidth to run any multimedia contents of even full-HD quality or beyond. Conversely, the WAN storage can assure

a plenty of storage space, whereas its I/O bandwidth is too low to play such a high-quality multimedia content. We found that WAN storage could provide an I/O bandwidth of 1MB/s in our experimental setup.

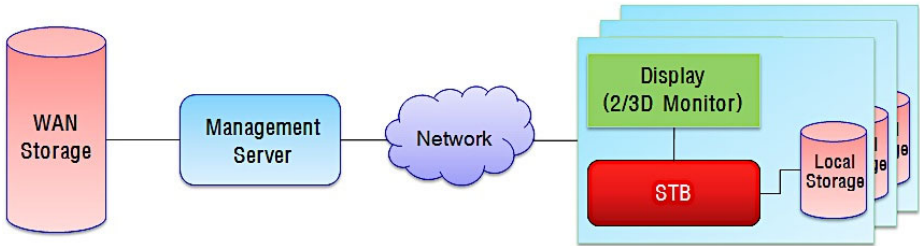


Fig. 1. Typical architecture of a networked digital signage system

Even though multimedia contents of different qualities from SD to full HD, we observed that the full-HD multimedia contents have been widely used for digital signage. In other words, there is no reason to use lower quality multimedia contents to capitalize on their products. Digital signage transmits a list of multimedia contents on a weekly basis or a monthly basis. On a monthly basis, we could see that approximately 1,000 contents are downloaded and being played at the digital signage by considering that 30 to 40 contents are in use per day. Assuming that transmitted contents are better than HD qualities, the total size to store all the contents becomes 1TB that is far beyond the maximum size of the local disk (200GB). On the contrary, the WAN storage can accommodate approximately three-month long multimedia contents (3,000-4,000 contents). Note that the WAN storage can not only store all the multimedia contents that the digital signage will download and play back without any size limitation that the local disk exhibits. However, it poses a critical I/O bandwidth problem. For example, downloading a normal 9-minute multimedia content of HD quality requires 20 minutes considering the WAN I/O bandwidth of 1MB/s.

In the meantime, iSCSI protocol-based storage (briefly iSCSI storage) has recently become prevalent as a different breed of storage architectures that transmits the traditional SCSI protocol over TCP/IP network [4][5]. Any host that accesses the iSCSI storage should have an iSCSI imitator driver (called iSCSI client) to send a SCSI command to the iSCSI storage over TCP/IP network. The iSCSI storage has an iSCSI target driver. We found that the iSCSI storage can provide higher than 2MB/s with light network traffic. However, this cannot be always guaranteed when considering ever-changing network traffic where the digital signage is connected. By knowing that a multimedia content of full-HD quality requires less than 2MB/s bandwidth, plugging the iSCSI storage into the current storage hierarchy for the digital signage (local disk on top of WAN storage) can be very attractive. Moreover, when using the iSCSI storage, we could easily remove the size limitation like the WAN storage. Thus, in this paper, we propose iSCSI-protocol based hybrid storage architecture for digital signage that can speed up the contents download time by inserting the iSCSI storage between the local disk and the WAN storage and by adding a novel content caching algorithm.

2 Hybrid Storage Architecture

2.1 Overall Architecture

Figure 2 shows an overall architecture of the proposed hybrid storage that contains a local disk, iSCSI storage, and WAN storage in a hierarchical manner and provides a content caching algorithm. Observe that the iSCSI storage sits between the local disk and the WAN storage. The digital signage media player decodes and outputs the multimedia contents from the local storage and Hstor-STR storage. The local storage has a limited size of 160GB, which stores multimedia contents that includes high quality contents like 1080p Full-HD. The content caching algorithm governs placements of multimedia contents inside the storage hierarchy, mostly for the local disk and the iSCSI storage. Notice that the iSCSI storage is divided into two distinctive regions: Hstor-STR and Hstor-NOSTR. The Hstor-STR contains any multimedia contents that can be directly played from the iSCSI storage without copying (downloading) them into the local disk. The Hstor-NOSTR contains other multimedia contents. The location of multimedia content can be determined dynamically depending on a demanded I/O bandwidth from the content and a current I/O bandwidth offered by the iSCSI storage.

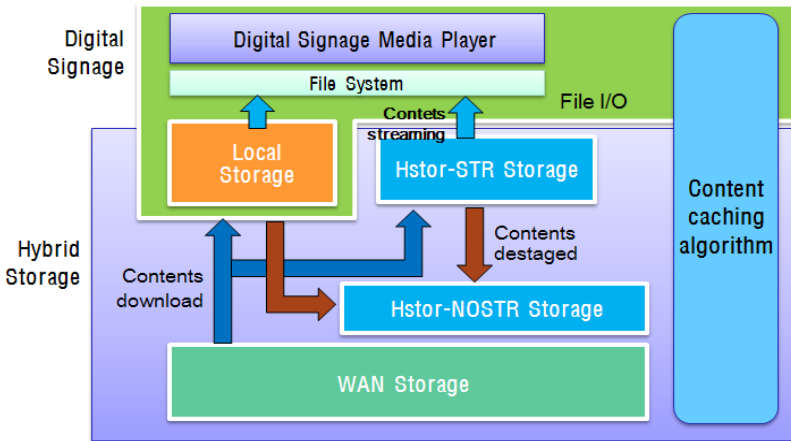


Fig. 2. The proposed iSCSI protocol based hybrid storage architecture

To begin with, it is assumed that all the multimedia contents are stored in the WAN storage. Given a playlist of multimedia contents of various qualities, the content caching algorithm works to have all the contents at either the local disk or the Hstor-STR. All the multimedia contents can be stored into the local disk, whereas some multimedia contents can be contained in the Hstor-STR that demands lower I/O bandwidth than that of the iSCSI storage. Note that the Hstor-NOSTR is regarded as a secondary cache to store any multimedia contents evicted from the local disk or the Hstor-STR. The detailed interactions among storage components will be given below.

2.2 Content Caching Algorithm

Algorithm 1 describes the proposed content caching algorithm working with the local disk, the two-iSCSI storage regions (Hstor-STR and Hstor-NOSTR), and the WAN storage. The key of the algorithm is to keep the frequently used multimedia contents as close as possible to the digital signage, eventually reducing its contents download time. The algorithm starts by examining if a multimedia content of the given playlist is existent in the local disk or Hstor-STR (line 1). If it is found, it can be reused after updating its associative LRU list (lines 2-3). Otherwise, it continues another search for the secondary cache (Hstor-NOSTR) (lines 5-6). The content found in the Hstor-NOSTR is copied (not moved) into either the local disk or the Hstor-STR (lines 7-15). The decision is made based on the aggregated LRU list for the two storage components. Otherwise, the content will be downloaded from the WAN storage. Depending on the relationship between its demanded I/O bandwidth and the I/O bandwidth of the iSCSI storage (line 7), the content can be stored into the Hstor-STR or the local disk. In either case, the victim contents of the content size should be chosen and then evicted from the storage.

Algorithm 1. Content caching algorithm for the hybrid storage

```

1:  foreach multimedia content in the playlist do
2:    if (the content is found in local disk) update its lru list
3:    else if (it is found in Hstor-STR) update its lru list
4:    else
5:      if (it is found in Hstor-NOSTR) update its lru list
6:      src is set to Hstor-NOSTR
7:      if (demanded I/O bandwidth of content <= Hstor-STR's I/O bandwidth)
8:        select victim contents from Hstor-STR
9:        destage(move) the victims to Hstor-NOSTR (using LRU policy)
10:       copy from Hstor-STR (or WAN) to Hstor-STR or local disk
11:     else
12:       select victims contents from local disk
13:       destage(move) the victims to Hstor-NOSTR (using LRU policy)
14:       copy from Hstor-STR (or WAN) to local disk
15:     endif
16:   endif
17: done

```

3 Performance Evaluation

3.1 Experimental Environment

We perform performance evaluations on a simulator and an actual digital signage (prototype). For both environments, the digital signage under test is configured with a 200GB local disk, 1TB iSCSI storage, WAN storage of unlimited size. The prototype is developed based on Sigma Design's SMP8654 [6]. SMP8654 is a embedded board

based on MIPS. It uses 500MHz CPU, 512MB DDR2 SDRAM, and is specialized playing multimedia. We use various workloads (playlists) that mix multimedia contents of different qualities (SD, HD and full-HD) evenly. We mainly measure the required time to have all the contents for a given playlist (called contents download time). In the experiment, the WAN storage is initially assumed to contain around 300 multimedia contents (used for a week, approximately 400GB in size). In workloads (playlists), we enforce different levels of locality in contents usage (locality of contents usage). For example, if the locality is 10%, 10% of the total contents will be used for playlists. Our evaluation considers the following three architectures: (1) the local disk only (LD), (2) a combination of the local disk and the iSCSI storage without the content caching algorithm (LD-iSCSI); (3) our proposed architecture (LD-iSCSI(STR, NO-STR)).

3.2 Performance Analysis

Figure 3 shows the results of performance gains (how much the contents download time of the given architecture can be reduced against the case of always downloading the entire contents of the playlist) with different locality of contents usage. In the experiments, it is configured that the I/O bandwidth of the iSCSI storage is 2MB/s and ratios of Hstor-STR to Hstor-NOSTR are configured 30% and 70%, respectively. Observe that the proposed architecture outperforms the other architectures maximally 30%. As expected, as the locality increases, the gain of the proposed architecture becomes much higher than the other architectures.

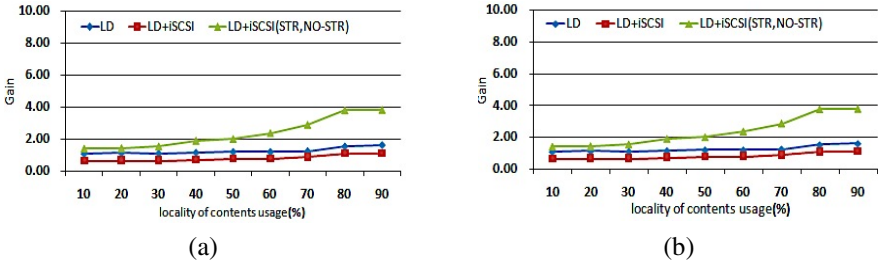


Fig. 3. Comparison of performance gains (against the case of always downloading the entire playlist) with different locality of contents usage in a simulator, where iSCSI storage = 2MB/s, a ratio of Hstor-STR to Hstor-NOSTR = (a) 30% and (b) 70%

Figure 4 shows the results of performance gains with different locality of contents usage in an actual digital signage (prototype). In the experiment, we found similar results in its prototype as in its simulation. There was no performance gain in the simulation when the locality of content usage becomes over 80% (Figure 3). However, observe that the performance gain continues to increase (Figure 4). Similarly, the overall performance gains of the proposed architecture are much higher than the other architectures.

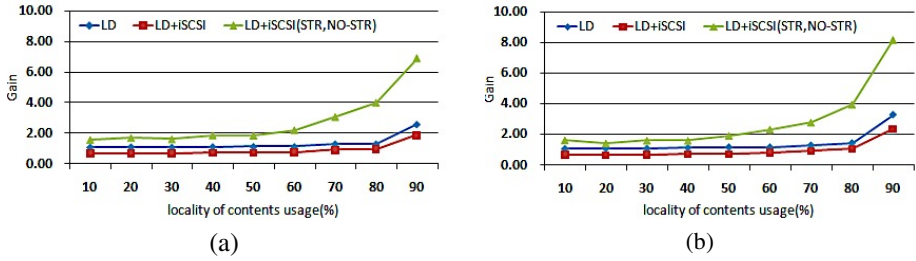


Fig. 4. Comparison of performance gains (against the case of always downloading the entire playlist) with different locality of contents usage in an actual digital signage (prototype), where iSCSI storage = 2MB/s, a ratio of Hstor-STR to Hstor-NOSTR = (a) 30% and (b) 70%

Figure 5 and 6 show the results of the performance gains with different iSCSI I/O speeds (bandwidth). Notice that no performance benefit can be obtained by using the proposed architecture when the iSCSI storage I/O speed becomes lower than 2MB/s, which is not enough performance bandwidth to play (stream) full-HD multimedia contents directly from the iSCSI storage. Once the iSCSI storage I/O speed reaches at 2MB/s, it does not affect the performance any more. The experimental results from the simulator and the prototype showed the same tendency. Besides, we can observe that the proposed architecture constantly outperforms the other architectures (on average two times better than the others in terms of the performance gain when its locality is 70%).

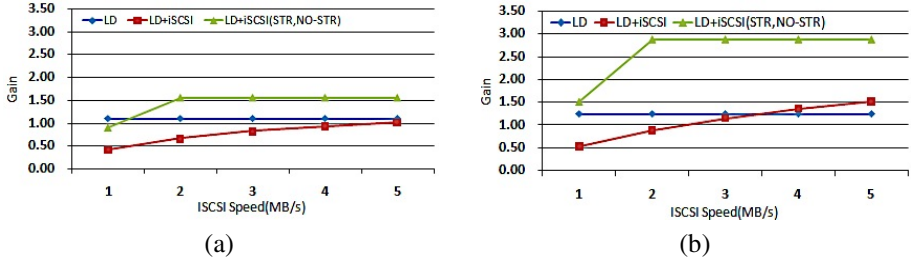


Fig. 5. Comparison of performance gains with different iSCSI storage I/O bandwidth in a simulator, where a ratio of Hstor-STR = 30%, locality of contents usage = (a) 30% and (b) 70%

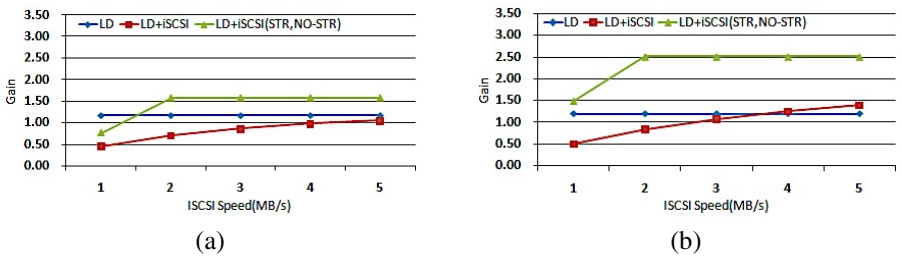


Fig. 6. Comparison of performance gains using different iSCSI storage I/O speed for each architecture in an actual digital signage (prototype), where a ratio of Hstor-STR = 30%, locality of contents usage = (a) 30% and (b) 70%

4 Concluding Remarks

This paper originally proposed the hybrid storage architecture for networked digital signage that additionally includes an iSCSI-proposal based storage layer and provides a novel content caching algorithm. Throughout extensive experiments on a simulator and its prototype, we found that the proposed hybrid storage architecture could reduce the contents download time at maximum 3 times than the other architectures.

Acknowledgement. This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)(NIPA-2011-C1090-1131-0009).

References

1. Maeda, K., Nishi, M., Yoshida, T., Suzuki, K., Inoue, H.: Digital Signage with Audience Detection Using TV Broadcasting Waves. In: 10th Annual International Symposium on Applications and the Internet, pp. 225–228 (2010)
2. Withana, A., Dayarathna, M., Inakage, M., Sugiura, K.: 2nd Design and Implementation of Distributed and Scalable Multimedia Signage System. In: Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, pp. 1–6 (2011)
3. Suzuki, K., Udaka, H., Yoshida, A., Nagao, J., Taniguchi, Y., Akutsu, A.: Metadata Delivery and Integration and Metadata Matching Technologies for Digital Signage. NTT Technical Review (2009)
4. Zhou, W.: Optimizing Institute network storage building base on iSCSI for IP-SAN. JCSSE, 98–100 (2010)
5. Singh, K., Singh, V., Lim, H., Cha, K.-H.: A Study on the Data Shaing in iSCSI Based Virtual Storage System. Korea Multimedia Society, 261–264 (2006)
6. Sigma designs,
http://www.sigmadesigns.com/uploads/documents/IPTV-8654_br.pdf

Adoption of Technology Applications in Healthcare: The Influence of Attitude toward Knowledge Sharing on Technology Acceptance in a Hospital

Pouyan Esmailzadeh, Murali Sambasivan, Naresh Kumar, and Hossein Nezakhati

Graduate School of Management, Universiti Putra Malaysia (UPM),
43400 UPM Serdang, Selangor, Malaysia
pouyanes@yahoo.com

Abstract. The use of different forms of technology has increased in healthcare profession recently. Clinical IT can change the practice patterns of healthcare professionals to improve the quality of health care delivery. When a new clinical IT is introduced in a hospital, healthcare professionals play an important role in the adoption and implementation process. But underutilization of clinical IT has emerged as a new challenge for the healthcare industry. So that healthcare professionals have not fully adopted the clinical IT systems. To improve overall acceptance of clinical IT in a hospital setting, this study (as a conceptual research) argues that unique feature of clinical IT can potentially affect healthcare professionals' adoption of new clinical IT system. This study proposes a modified technology adoption model (TAM) to incorporate both the special characteristic of healthcare professionals and unique feature of clinical IT. This study discusses that how attitude toward knowledge sharing affects healthcare professionals' intention to use the clinical IT system.

Keywords: Clinical IT, TAM, Perceived threat to professional autonomy, Attitude toward knowledge sharing, Healthcare professionals' involvement in decision making, Perceived usefulness, Perceived ease of use.

1 Introduction

User adoption of new technology has attracted the attention of a large number of researchers in Information Systems (IS) studies. As technology innovation can leave its mark on improving the productivity and performance (at different levels) of organizations [1], researchers have been trying to find out factors affecting the successful adoption of technological advancement among users. One of the most important determinants influencing the success of Information Technology (IT) adoption is that to what degree IT systems are accepted by users [2]. User acceptance is defined as the willingness of the users to use IT which is designed to support tasks [3]. Organizations invest in IT systems with the hope of cutting costs, increasing the quality of products or services [4]. But if users do not accept the systems, the organizations can not benefit significantly from the new systems. On the other hand, if users accept new IT systems they become more willing to make use of the new

systems [5]. The usage of a newly introduced IT system can be a sign of the IT system success [6]. Therefore, finding the reasons that motivate people to use the new systems or understanding the source of resistance toward using new systems is important to both system designers and developers [7].

The use of IT in health care practices has increased recently [8]. A variety of IT systems such as clinical information systems, personal digital assistants, electronic patient records and other applications have gradually become established in the healthcare industry. Clinical IT applications in healthcare are regarded as a key element in improving the quality of medical care. However, factors affecting healthcare professionals' adoption behavior regarding the IT systems is not completely clear yet ([9],[10],[11]). The concern of having new clinical IT systems unused is still one of the biggest issues for the clinical IT developers ([2], [12]).

With reference to Walter and Lopez [11], two types of IT are available in medical care environment. The first one is Electronic Medical Records (EMR) systems which are computer systems that allow users to create, store, and retrieve patient charts on a computer. The second one is Clinical Decision Support (CDS) system that is classified as a decision support system. A CDS System is regarded as an application of Decision Support System (DSS), which takes patient data as input and generates decision- specific advice ([13], [14]). These systems are referred to as knowledge-based systems that use patient data and series of reasoning techniques to generate diagnostic and treatment options and care planning.

When a new clinical IT is introduced in a hospital, healthcare professionals play an important role in the adoption and implementation process. Thus, healthcare professionals need to use emerging clinical IT to reap the benefits of new systems otherwise the technology will remain underutilized. In other words, healthcare professional's acceptance is reported as an important need to the success of the clinical systems ([15], [2]). Once the users accept new IT, they are more likely to make changes in their existing work routines and incorporate the new IT into the flow of their everyday work practices [11].

According to King and He [16], in recent years there has been an increasing interest in the identification of factors that cause potential users to accept and take advantage of systems developed and implemented by others. With respect to individual intention to accept new technology, several studies have been conducted and eight theoretical models have been developed: Theory of Reasoned Action (TRA), Technology Acceptance Model (TAM), Motivational Model (MM), Theory of Planned Behavior (TPB), a combined Theory of Planned Behavior/Technology Acceptance Model (C-TAM-TPB), Model of PC Utilization (MPCU), Innovation Diffusion Theory (IDT), and Social Cognitive Theory (SCT). Each model explains the user's individual readiness to accept new information systems and technology [17].

Despite the large volume of studies in technology acceptance research, very little work has been done in the health care context [9]. This is a sign of a significant gap in this area. Also, based on the studies conducted in health sector, it is shown that healthcare professionals have not fully adopted the new IT systems ([10],[11], [18], [19] [20], [21], [22],[23]).

Although TAM has been used as a useful tool to explain technology acceptance process in different fields, this model is not applicable for healthcare professionals

[19]. For instance, this model is still very general and is not designed for any particular profession [11]. Each profession has special contextual characteristics that may affect IT adoption behavior. For instance, unique characteristics of IT users should be included in IT adoption models in order to better explain their intention to accept new technology.

Little research has studied special characteristics of healthcare professionals and different features of clinical information systems in the field of healthcare professional adoption [11]. Furthermore, according to Moon and Kim [24], besides the important constructs embedded in technology adoption models, additional exploratory factors are needed to better account for the variance of accepting new technology such as the specific technology context.

2 Theoretical Background

2.1 Technology Acceptance Model (TAM)

Based on a body of literature, TAM is the most influential IT adoption model and is widely applied to explain the IT acceptance process in different contexts [25]. Davis developed TAM based on TRA in 1989 to mainly explain technology use in various situations and cultures in order to increase user acceptance of systems. Another reason for usefulness as well as popularity of TAM is its parsimony, simplicity, understandability and gaining empirical support within a variety of user groups [26].

The original TAM suggests that two beliefs namely, perceived usefulness and perceived ease of use play a pivotal role in underscoring individual acceptance of a new technology [27,28]. The first variable, perceived usefulness, is considered as the degree to which a person believes that by using a particular system his/her job performance would be enhanced [29]. The second one, perceived ease of use, is operationally defined as the extent to which a person believes that using a particular system would be effortless [29]. These factors can be addressed during the system development stage to solve the users' acceptance problem [30]. These factors determine behavioral intention that is found by a wide number of studies [31], as a better predictor of actual system usage. In the field of social science, intention to use a new IT is defined as user willingness to actual behavior of using the new IT.

2.2 Theory of Professionals

According to Sharma [32], the holders of some occupations (such as medical practice) are defined professionals. The healthcare professionals who considered in this study consist of all kind of physicians and specialists from different medical specialty areas. This group can make use of clinical IT potentials to improve health care delivery and efficiency. Professionals have been attributed some unique and professional characteristics that make them different from other non-professionals. As stated by Brennan and Coles [33], healthcare professionals' professionalism is rooted in a set of values. The most important characteristic is healthcare professional autonomy and the other features are patient sovereignty, physician confidentiality, and habits of learning. According to Chau and Hu [34], the differences between healthcare

professionals and other user groups in terms of accepting new IT derive from a set of values such as:

- Specialized training which is obtained over a considerable period gives them the knowledge and expertise that is required in this profession [35].
- Professional autonomy is defined as the control that professionals have over the processes, conditions and content of their medical practice [36]. Literature states that professional autonomy is the most important professional value provided for healthcare professionals [37].
- The third characteristic is professional work arrangements where healthcare professionals are considered as health care providers, hospitals become health care facilities, and patient are both the product and the client in the healthcare environment [35, 38]. Beside professionals, there are two other occupational groups working in a hospital. The para-professional group, such as medical assistants, owns only partial professional knowledge and skill and assists healthcare professionals in their healthcare practices. The last group is non-professionals who are just prepared to engage in running clerical, office work and administrative duties.

Due to some privileges originate from professional autonomy, healthcare professionals have power over non-professionals and para-professionals and can control the tasks conducted by them [39]. Therefore, healthcare professionals try to support the factors that strengthen their professional autonomy and resist the factors that may erode their professional autonomy [11]. Despite the significant role of professional autonomy in healthcare professionals IT adoption behavior, less emphasis has been placed to explore whether and how this central characteristic influences healthcare professional's acceptance of new clinical IT [11].

Adams [40] has argued that previous studies highlighted some similar constructs like behavioral control and professional autonomy has not been studied as a central characteristic of healthcare professionals. Walter and Lopez [11], first introduced perceived threat to professional autonomy as a new construct in studying IT acceptance. This construct is operationally defined as "the degree to which a person believes that using a particular system would decrease his/her control over the conditions, processes, procedures, or content of his/her work".

2.3 Factor Affecting Perceived Threat to Professional Autonomy

A considerable amount of literature has been published to show that the autonomous practice of physicians and their independence in decision making are two distinct characteristics of medical profession [41]. In addition, like other professionals, healthcare professionals are highly committed to their own profession and their performance is evaluated based on a subjective peer review process. Yet, healthcare professionals believe that their professional autonomy is influenced by a technology that is supposed to change physicians' long-established practice pattern.

According to a rich body of literature, a feature of clinical IT can be considered as threatening factors to healthcare professional autonomy [11]. The factor is the level of knowledge codification and knowledge distribution conducted by a clinical IT system. Due to the possession of abstract and expert body of medical knowledge by healthcare professionals, they are less likely to accept the type of IT system that organize, codify

and distribute their knowledge which mainly makes them distinct from other occupational groups.

According to Walter and Lopez [11] different types of clinical IT (EMR and CDS), to some extent, involve knowledge codification and thus make physicians' analytical views and expert knowledge accessible to the subordinate group (such as non-professionals and paraprofessionals). Based on the existing literature on knowledge management, knowledge codification refers to converting tacit knowledge into explicit knowledge in a way that it can be usable by all the organizational members [42]. Therefore, knowledge codification will lead to more knowledge distribution and contributes to knowledge sharing in organizations.

According to Prasad and Prasad [43], the practice of knowledge codification enables para-professionals to get access to greater amount of knowledge in organizations. Therefore, more knowledge codification leads to more knowledge distribution in a hospital. Harrison et al. [44] have mentioned that knowledge codification abolishes the exclusive state of having abstract, expert and unshared medical knowledge, as well as the sole rights of having specialized competence and expertise possessed by healthcare professionals.

Moreover, measures for healthcare professionals' performance evaluation become more objective (rather than subjective) when their esoteric knowledge becomes more easily reachable to the subordinate group. Also due to the implementation of a clinical IT, the physicians' diagnostic decisions and any part of patient care practice is easily monitored by others such as physician assistants, paraprofessionals and non-professionals working in the health care environment and considered as the physicians' subordinates. Therefore, being monitored by others (especially from out of the profession) may increase the possibility of reviewing the physicians' prescriptions and revealing their patient care process. Consequently, monitoring the treatment options prescribed by physicians can intensify perceived threat to professional autonomy.

As a result, healthcare professionals can no longer claim exclusive possession of a body of specialized knowledge and might not have control over resources and the tasks done by the subordinate group. Kritzer [45] has supported the idea that exclusive ownership of esoteric body of knowledge leads to professional autonomy. This study has also carried out an investigation into the relationship between the exposure of physicians' abstract knowledge to subordinate groups and decreasing their professional autonomy. Therefore, drawing on the above arguments, perceived threat to professional autonomy is mainly intensified through the increased level of knowledge codification by means of clinical IT system. McLaughlin and Webster [46] have published a case study in which they described the effect of IT on professional autonomy. They conclude that physicians' professional autonomy is reduced as their abstract knowledge is codified by the IT system. Also, they have stated that IT system can blur occupational boundaries among different occupational groups.

To sum up, healthcare professionals believe that clinical IT can codify their esoteric knowledge to a high extent and in turn distribute their knowledge to all physician assistants, non-professionals and para-professionals in a hospital setting. By doing so, healthcare professionals can no longer claim on having the abstract knowledge and they will lose their control over the tasks performed by the subordinate group. Therefore, if the features of a clinical IT contributes to knowledge codification and knowledge

sharing, more threat will be perceived by healthcare professionals to their professional autonomy and they become less likely to use the clinical IT system.

2.4 Knowledge Sharing in a Hospital Setting

Healthcare professionals are seen as knowledge-intensive employees working in a hospital. Due to the professional autonomy, physicians usually do not consider the subordinate group as their co-workers and they are less likely to share their knowledge with them. But it should be noted that other occupational groups working in a hospital (such as physician assistants and para-professionals) can help them better complete the process of disease diagnosis and treatments. The role of the subordinate group becomes more significant when knowledge sharing environment is built between professionals and this group. An issue on knowledge sharing between professionals and the subordinate group is related to the nature of knowledge sharing between subordinates and supervisors which is usually bounded by formal relationships [47]. Hence, healthcare professionals do not want their subordinate group (such as physician assistants and para-professionals) to access to their knowledge by using the clinical IT in order to maintain their professional autonomy.

However, according to Dexter et al. [48], in some cases it has been reported that CDS becomes more effective if other clinicians (such as nurses) receive and use the information delivered by the system. As a conclusion, if healthcare professionals share the required knowledge and expertise with the subordinate group, this group can work effectively and also they can better assist healthcare professionals in the competition of delivering health services.

To share knowledge with a hospital's members two issues are essential. First, ideas and insights should be presented in acceptable and understandable forms with the intention that they can be clearly received, interpreted and used by others. Second, individuals with ideas should be eager to practice knowledge sharing for the benefit of the entire organization [49].

Healthcare professionals' attitude toward knowledge sharing can improve the exchange of ideas and in turn collaboration with other occupational groups. Although healthcare professionals have command over medical knowledge, the most reliable personnel in terms of serving care-giving and nursing practices are para-professionals. Thus, healthcare professionals should consider other occupational groups as a team member and they can enhance the quality of treatment given to patients through a reciprocal relationship and collaboration [50]. As a result, if healthcare professionals have a positive attitude toward knowledge sharing with physician assistants and para-professionals and they perceive that shared knowledge is effective for their organizations (hospitals). In turn, it may reduce the threats they perceive from clinical IT on distribution of their knowledge to the subordinate group.

2.5 Level of Healthcare Professionals' Involvement in Decision Making

Literature indicates that one of the most important concerns in the development and implementation of an IT system is related to managerial point of view. According to a suggestion made by Hoslinger and Beaton [50], instead of requesting healthcare professionals to accomplish more in compressed schedules they should be asked to

manage the resources. Managing time and data with the aid of health information technology can help them to be successful in offering more tailored healthcare and diagnosis to increased number of patients. In a case study conducted by Kohli et al. [51], on the subject of hospital-physician collaboration, it is found that success of an IT project can be guaranteed with the help of a precise strategy to involve physicians both in planning and implementation. Further, trust between physicians and the hospital is found to be an important requirement contributing to changes in clinical practices on the hospital floors. It is reported that even well-designed information systems cannot succeed to show the conversion effectiveness due to the lack of system use or the failure of the users to make changes when is needed [51].

As a conclusion, according to the qualitative study by Kohli et al.[51], healthcare professionals are more likely to change their day-to-day work activities when they have the right of being involved in both planning and implementation of a web-based physician profiling system and such willingness leads to the success of this system. Moreover, Walter and Lopez [11] have suggested that although physicians are very vulnerable and reactive to the type of clinical IT facilitating knowledge dissemination, such concerns may be reduced by the degree of being involved in making relative decisions about the development of that IT.

3 Conceptual Framework and Hypotheses Development

In the proposed framework, the dependent variable is healthcare professionals' intention to use the clinical IT. The construct perceived threat to professional autonomy indicates the threat observed by healthcare professionals on their central characteristic that is autonomous practice. The factor affects this construct stems from special feature of the clinical IT. The factor is the level of knowledge codification and sharing with subordinate group (such as para-professionals and physician assistants). This factor provides a human-oriented stream. In human-oriented stream, healthcare professionals' attitude toward knowledge sharing is defined as the main construct. The following figure (Fig.1) depicts the conceptual framework of this study.

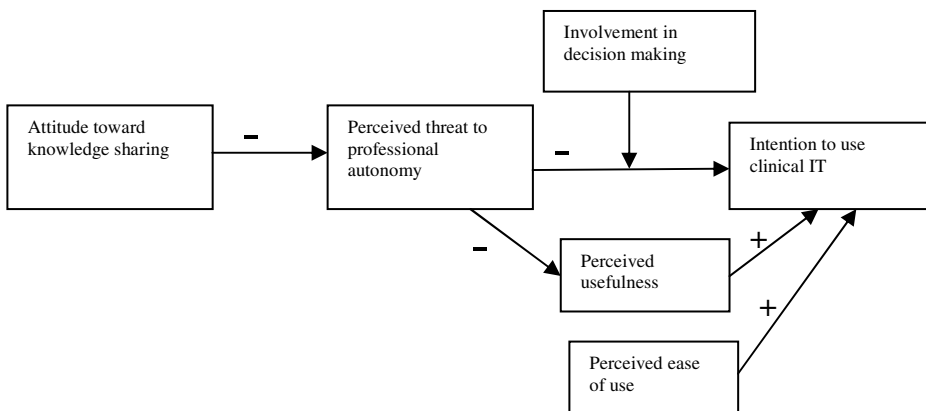


Fig. 1. Proposal framework

3.1 Perceived Threat to Professional Autonomy and Intention to Use Clinical IT

Intention to use new technology is the dependent variable which refers to individual intention or readiness to accept a new technology [29]. Also according to Walter and Lopez [11], perceived threat to professional autonomy is defined as “the degree to which a person believes that using a particular system would decrease his or her control over the conditions, processes, procedures, or content of his or her work”. This study proposes that perceived threat to professional autonomy reduces healthcare professional intention to use clinical IT in a hospital setting. It implies that if healthcare professionals perceive the application of clinical IT as threatening to their professional autonomy, the possibility of using clinical IT by them will decrease. Therefore, the first proposition is developed as follows:

Proposition 1. *Healthcare professionals’ intention to use clinical IT in a hospital setting is negatively related to perceived threat to professional autonomy.*

3.2 Perceived Usefulness and Intention to Use Clinical IT

Consistent with prior studies in the healthcare context, Yi et al. [18] have stated that perceived improved performance resulting from using IT strongly affects physicians’ intention to use the system in the healthcare sector. The significant role of perceived usefulness among physicians in shaping their intention toward using a new technology might have been centered on physicians’ utility-based point of view about using technology [34]. It means they accept a new technology when it possesses desired utility and becomes instrumental in their practices. According to Chang et al. [14], perceived usefulness exerts the most significant impact on physicians’ intention to use CDS. Based on Kijisanayotin et al. [2], perceived usefulness is the most important determinant of intention to use health IT in a developing countries healthcare context. Therefore, as long as healthcare professionals perceive clinical IT as a source of performance enhancement, they become more willing to use the system. Thus, the next proposition is developed as follows:

Proposition 2. *Healthcare professionals’ intention to use clinical IT in a hospital setting is positively related to their perceived usefulness.*

3.3 Perceived Threat to Professional Autonomy and Perceived Usefulness

As mentioned by Walter and Lopez [11], perceived usefulness gained from using an IT is not always what physicians are concerned about and if the clinical IT system erodes their professional autonomy performance expectancy becomes insignificant for healthcare professionals. In the healthcare context, if a clinical IT invalidates healthcare professional autonomy and changes their practice patterns, the system would not be fully used for the purpose of performance improvement expected from the clinical IT. This may occur to maintain professional autonomy instead of the new useful clinical IT that is perceived as threatening to their professional autonomy. Thus, the next proposition manifests this effect as follows:

Proposition 3. *Healthcare professionals’ perceived usefulness is negatively related to perceived threat to professional autonomy in a hospital setting.*

3.4 Perceived Ease of Use and Intention to Use Clinical IT

In line with Davis [29], intention to use new IT systems is positively related to perceived ease of use. Chang et al. [14] have stated that effort expectancy is a significant predictor for physicians' intention to use CDS. As supported by Kijisanayotin et al. [2], effort expectancy is a key factor in shaping physicians' intention to use technology. Therefore, if healthcare professionals find the new clinical IT easy to understand and use, they become more likely to use the system in their practice pattern. Thus, the next proposition states this idea as follows:

Proposition 4. *Intention to use clinical IT is positively related to perceived usefulness in a hospital setting.*

3.5 The Moderating Effect of Healthcare Professionals' Level of Involvement in Decision Making

Hoslinger and Beaton [50] have considered managerial roles for healthcare professionals. The literature states that if healthcare professionals involve in decision making process regarding the introduction and development of appropriate IT system in organizations, the threat perceived from clinical IT can be reduced and they become more willing to change their long-term practice pattern [51]. Therefore, this study proposes that healthcare professional's high level of involvement in decision making regarding the development and implementation of clinical IT can reduce the negative effect of threat they perceive from the system and consequently may improve their intention to use the system. Thus, the following proposition is developed as:

Proposition 5. *Healthcare professionals' level of involvement in decision making about the clinical IT moderates the relationship between perceived threat to professional autonomy and intention to use the system.*

3.6 Healthcare Professionals' Attitude toward Sharing Knowledge and Perceived Threat to Professional Autonomy

According to King [16], knowledge sharing occurs by distribution of knowledge through a system (repository) with individuals who are not usually familiar to the contributor. In this case, healthcare professionals' knowledge is shared by the clinical IT systems with the subordinate group (such as para-professionals, physician assistants and junior healthcare professionals). But in the light of claiming exclusive possession of esoteric bodies of medical knowledge, healthcare professionals are less likely to use this system. They perceive that this type of clinical IT can weaken their professional autonomy.

According to the new definition of professionalism suggested by Holsinger and Beaton [50], corporate values (professionalism and teamwork) should be applied in a hospital setting. It means collaboration and teamwork should be established within healthcare professionals and also between healthcare professionals and other occupational groups. Knowledge sharing is a sign of collaboration and inter-organizational relationship between occupational groups in a hospital [52]. Thus, once

healthcare professionals hold a positive attitude toward knowledge sharing and they perceive that shared knowledge is effective for their organization (hospital), they feel less threatened by the clinical IT system.

Healthcare professional's attitude toward knowledge sharing has been defined as the degree to which they have a favorable or unfavorable evaluation of knowledge sharing ([53],[54],[55], [56],[57],[58]). As mentioned by Kwok and Gao [59], an individual's attitude toward knowledge sharing can affect his/her intention to share knowledge and in turn influence actual behavior. As a result, one way to reduce the negative effects of perceived threat to healthcare professionals' autonomy can be associated with their attitude toward knowledge sharing in hospitals. When healthcare professionals hold a favorable attitude toward knowledge sharing in a hospital setting, they may perceive less threat by clinical IT that is supposed to distribute their knowledge among other organizational members. Consequently, they may become more likely to use the clinical IT system. Therefore, the related proposition is developed as follows:

Proposition 6. *There is a negative relationship between healthcare professionals' attitude toward knowledge sharing and the perceived threat to professional autonomy.*

4 Implications of the Study

4.1 Theoretical

From a theoretical standpoint and theory building, the research contributes to IT adoption theories explaining healthcare professionals' intention to accept new technology. Since the TAM is general and cannot address healthcare professionals' unique characteristics, this model has been improved to fit the healthcare context and better explain healthcare professionals' IT adoption behavior in a hospital setting. The influential constructs from the medical literature are perceived threat to professional autonomy and healthcare professionals' involvement in decision making in hospitals. The construct healthcare professional attitude toward knowledge sharing is borrowed from knowledge management related literature. To improve the overall healthcare professionals' IT acceptance, this study proposes to improve healthcare professionals' attitude toward knowledge sharing in hospitals. Besides, this study argues that level of healthcare professionals' involvement in decision making can reduce threat perceived from the clinical IT and improve their intention to use the IT system. In other words, a modified model is proposed to explain and predict use of new technology in the healthcare industry.

4.2 Practical

From a practical point of view, this study proposes that hospital management should devise a strategy to improve healthcare professionals' attitude toward knowledge sharing in hospitals. To do so, hospital management may focus on the development of

a social network and shared goals between healthcare professionals and the subordinate group. Besides, healthcare management may improve the level of healthcare professionals' involvement in decision making on development and implementation of the clinical IT systems. With this understanding hospital management can reduce the negative effects of perceived threat to professional autonomy and improve overall acceptance of clinical IT by healthcare professionals.

References

1. Orlikowski, W.J.: Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science* 11(4), 404–428 (2000)
2. Kijsanayotin, B., Pannarunothai, S., Speedie, S.M.: Factors influencing health information technology adoption in Thailand's community health centers: Applying the UTAUT model. *International Journal Medical Informatics* 79, 404–416 (2009)
3. Dillon, A., Morris, M.: User acceptance of information technology: theories and models. In: Williams, M. (ed.) *Annual Review of Information Science and Technology*, vol. 31, pp. 3–32 (1996)
4. Lederer, A.L., Maupin, D.J., Sena, M.P., Zhuang, Y.: The role of ease of use, usefulness and attitude in the prediction of World Wide Web usage. In: *Proceedings of the 1998 Association for Computing Machinery Special Interest Group on Computer Personnel Research* (1998)
5. Succi, M.J., Walter, Z.D.: Theory of user acceptance of information technologies: An examination of health care professionals. In: *32nd Hawaii International Conference on System Sciences*. IEEE Computer Society, Hawaii (1999)
6. Pikkarainen, T., Pikkarainen, K., Karjaluoto, H.: Consumer acceptance of online banking: an extension of the technology acceptance model. *Internet Research* 14(3), 224–235 (2004), <http://www.emeralinsight.com/researchregister>
7. Mathieson: Predicting user intention: comparing the technology acceptance model with theory of planned behavior. *Information Systems Research* 2(3), 173–191 (1991)
8. Obstfelder, A., Engeseth, K.H., Wynn, R.: Characteristics of successfully implemented telemedical applications. *Implement Sci.* 2(25) (2007)
9. Aggelidis, V.P., Chatzoglou, P.D.: Using a modified technology acceptance model in hospitals. *International Journal of Medical Informatics* 78(2), 115–126 (2009)
10. Tung, F.C., Chang, S.C., Chou, C.M.: An extension of trust and TAM model with IDT in the adoption of the electronic logistics information system in HIS in the medical industry. *Int. J. Med. Inform.* 77(5), 324–335 (2008)
11. Walter, Z., Lopez, M.S.: Physicians acceptance of information technology: Role of perceived threat to professional autonomy. *Decision Support Systems* 46(1), 206–215 (2008)
12. Gagnon, M.P., Pluye, P., Desmartis, M., Car, J., Pagliari, C., Labrecque, M., Fremont, P., Gagnon, J., Nojya, M., Legare, F.: A systematic review of interventions promoting clinical information retrieval technology (CIRT) adoption by healthcare professionals. *International Journal of Medical Informatics* 79, 669–680 (2010)
13. Van Bommel, J.H., Musen, M.A.: *Handbook of medical informatics*. Springer, NY (1997)
14. Chang, I.-C., Hwang, H.-G., Hung, W.-F., Li, Y.-C.: Physicians' acceptance of pharmacokinetics-based clinical decision support systems. *Expert Systems with Applications* 33(2), 296–303 (2007)

15. Payton, F.C.: Lesson learned from three interorganizational health care information systems. *Information and Management* 36(6), 311–322 (2000)
16. King, W.R., He, J.: A meta-analysis of the technology acceptance model. *Information and Management* 43, 740–755 (2006)
17. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology* 22(14), 1111–1132 (1992)
18. Yi, M.Y., Jackson, J.D., Park, J.S., Probst, J.C.: Understanding information technology acceptance by individual professionals: Toward an integrative view. *Information and Management* 43(3), 350–363 (2006)
19. Chismar, W.G., Wiley-Patton, S.: Does the extended technology acceptance model apply to physicians. In: *Proceedings of the 36th Hawaii International Conference on System Sciences* (2003)
20. Dearne, K.: Health's tech bypass. In: *The Australian*, Sydney, p. 1, 4 (2003)
21. Murray, D.: Healthcare challenge. *Australian Information Week* 3, 10–18 (2002)
22. Wenn, A., Tantnall, A., Sellitto, C., Darbyshire, P., Burgess, S.: A sociotechnical investigation of factors affecting IT adoption by rural GPs'. In: *Information Technology in Regional Areas. Using IT: Make IT happen*, Online, Rockhampton, Queensland Australia (2002)
23. Western, M., Dwan, K., Makkai, T., del Mar, C., Western, J.: *Measuring IT use in Australian General Practice*. University of Queensland, Australia (2001)
24. Moon, J.W., Kim, Y.G.: Extending the TAM for a world-wide-web context. *Information & Management* 38(4), 217–230 (2001)
25. Hossain, L., de Silva, A.: Exploring user acceptance of technology using social networks. *Journal of High Technology Research* 20(1), 1–18 (2009)
26. Wang, Y.S., Wang, Y.M., Lin, H.H., Tang, T.I.: Determinants of user acceptance of Internet banking: an empirical study. *International Journal of Service Industry Management* 14(5), 501–519 (2003)
27. Anandarajan, M., Igbaria, M., Anakwe, U.P.: IT acceptance in a less developed country: A motivational factor perspective. *International Journal of Information Management* 22(1), 47–65 (2002)
28. Ghorab, K.E.: The impact of technology acceptance considerations on system usage and adopted level of technological sophistication: An empirical investigation. *Journal of Information Management* 17(4), 249–259 (1997)
29. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3), 319–340 (1989)
30. Taylor, S., Todd, P.A.: Assessing IT Usage: The Role of Prior Experience. *MIS Quarterly* 19(2), 561–570 (1995)
31. Venkatesh, V., Davis, F.D.: A model of the perceived ease of use: Development and test. *Decision Sciences* 27, 451–481 (1996)
32. Sharma, A.: Professionals as agent: knowledge asymmetry in agency exchanges. *Academy of Management Review* 22(3), 758–798 (1997)
33. Brennan, M., Coles, C.: Developing professional skills. *The Lancet* 362(9394), 362–1506 (2003)
34. Chau, P.Y.K., Hu, P.J.: Examining a model of information technology acceptance by individual professionals: an exploratory study. *Journal of Management Information Systems* 18(4), 191–229 (2002)
35. Watts, C.: Erosion of healthcare professional autonomy and public respect for the profession. *Surgical Neurology* 71(3), 269–273 (2008)

36. Raelin, J.: An anatomy of autonomy: managing professionals. *The Academy of Management Executive* 3(3), 216–228 (1989)
37. Zuger, A.: Dissatisfaction with medical practice. *N. Engl. J. Med.* 350(1), 69–75 (2004)
38. Montague, E.N.H., Kleiner, B.M., Winchester, W.W.: Empirically understanding trust in medical technology. *International Journal of Industrial Ergonomics* 39(4), 628–634 (2009)
39. Freidson, E.: *Profession of Medicine: A Study of the Sociology of Applied Knowledge*. The University of Chicago Press, Chicago (1988)
40. Adams, D.W.: Standards and the development of professions: Implications for educational evaluation. In: Paper presented at the 64th Annual Meeting of the American Educational Research Association (1980), <http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/> (retrieved October 16, 2008)
41. Tanriverdi, H., Venkatraman, N.: Creation of professional networks: an emergent model using telemedicine as a case. In: 32nd Hawaii International Conference on System Sciences. IEEE Computer Society, Hawaii (1999)
42. Zack, M.H.: Managing codified knowledge. *Sloan Management Review* 40(4), 45–58 (1999)
43. Prasad, P., Prasad, A.: The ideology of professionalism and work computerization: institutionalist study of technological change. *Human Relations* 47(12), 1433–1458 (1994)
44. Harrison, S., Dowswell, G., Wright, J.: Practice nurses and clinical guidelines in a changing primary care context: an empirical study. *Journal of Advanced Nursing* 39(3), 299–307 (2002)
45. Kritzer, H.M.: The professions are dead, long live the Professions: Legal practice in a post professional world. *Law & Society Review* 33(3), 713–759 (1999)
46. Mclaughlin, J., Webster, A.: Rationalizing knowledge: IT systems, professional identities and power. *The Sociological Review* 46(4), 781–802 (1998)
47. Yang, S.C., Farn, C.K.: Social capital behavioral control, and tacit knowledge sharing- A multi-informant design. *International Journal of Information Management* 29(3), 210–218 (2009)
48. Dexter, P.R., Perkins, S.M., Maharry, K.S.: Inpatient computer-based standing orders vs. physician reminders to increase influenza and pneumococcal vaccination rates: a randomized trial. *JAMA* 292(19), 2366–2371 (2004)
49. Jabr, N.H.: Healthcare professionals' attitudes towards knowledge transfer and sharing. *International Business Journal* 17(4), 248–260 (2007)
50. Holsinger, J., Beaton, B.: Physician professionalism for a new century. *Clinical Autonomy* 19(5), 473–479 (2006)
51. Kohli, R., Piontek, F., Ellington, T., VanOsdol, T., Shepard, M., Brazel, G.: Managing customer relationships through E-business decision support applications: a case of hospital-physician collaboration. *Decision Support Systems* 32(2), 171–187 (2001)
52. Bartunek, J.M.: Intergroup relationships and quality improvement in healthcare. *BMJ Qual. Saf.* 20, 62–66 (2010)
53. Ajzen, I.: Constructing a TPB questionnaire: conceptual and methodological considerations (2001a), <http://www.unix.oit.umass.edu/tpb.htm> (Retrieved March 24, 2009)
54. Ajzen, I.: Nature and operation of attitudes. *Annales in Review of Psychology* 52, 27–58 (2001b)
55. Ajzen, I.: The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50(2), 179–211 (1991)

56. Bock, G.W., Kim, Y.G.: Breaking the myths of rewards: an exploratory study of attitudes about knowledge sharing. *Information Resources Management Journal* 15(2), 14–21 (2002)
57. Chang, M.K.: Predicting unethical behavior: a comparison of the theory of reasoned action and the theory of planned behavior. *Journal of Business Ethics* 17(16), 1825–1834 (1998)
58. Chau, P.Y.K., Hu, P.J.H.: Information technology acceptance by individual professionals: a model comparison approach. *Decision Sciences* 32(4), 699–719 (2001)
59. Kwok, S.H., Gao, S.: Attitude toward knowledge sharing behavior. *Journal of Information Systems* 46(2), 45–51 (2006)

Generic Evolutionary Framework for Simulating Business Processes

Veng-Ian Chan and Yain-Whar Si

Faculty of Science and Technology, University of Macau
{ma76505, fstasp}@umac.mo

Abstract. Business process simulation enables detail analysis of resource allocation strategies without actually deploying the processes. Although business process simulation has been widely researched in recent years, less attention has been devoted to automating the simulation of business processes with the help of evolutionary computation. In this research, we aim to implement a generic GA modeling framework which can be used to simulate different kinds of business processes. Specifically, optimum resource allocation scheme for the simulation can be effectively chosen by the evolution process of a genetic algorithm (GA). The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from a given business process and can produce the best resource assignment scheme.

Keywords: Genetic algorithm, resource allocation, business process, Color Petri Nets.

1 Introduction

Business Process Simulation (BPS) [8] artificially implements and assists the management of change in a variety of manufacturing and service settings. BPS enables detail analysis of such performance indicators without actually deploying the processes. BPS is widely used for analyzing business process models in many different domains. For instance, simulation was used to study the effect of stochastic customer shopping traffic for the IBM's Personal Computer Division [9] as well as for analyzing the impact on the user experience and the cost of using the application when a mobile channel presentation (based on GSM, HSCSD, GPRS and UMTS networks) is added [10].

Configuration of process models includes assigning tasks with available resources and inputting temporal constraints. Although BPS has been extensively used in performance analysis, configuration of process models for BPS has been mostly done manually by analysts. As a result, finding the right resource assignment scheme for simulation can be time consuming since it is usually done by trial-and-error approach. In addition, a large number of resource assignment schemes may exist for simulating a complex process model. With the help of evolutionary computation, optimum resource assignment for business process simulation can be effectively chosen by the

evolution process of a Genetic Algorithm (GA) [11]. A Genetic Algorithm can perform a powerful form of hill climbing search by maintaining multiple solutions and eliminating unpromising solutions.

Although GAs can be used to choose optimal resource assignment for BPS, experiments are only limited to modeling process models from a number of application scenarios. In these scenarios [12, 13], different GA designs are required for capturing process dependent parameters. For instance, chromosomes design, crossover and selection operators, and fitness functions are different for each case and they are required to be tailored for each underlying BPS problem. In this research, we propose a generic GA modeling framework which can be used to simulate different kinds of business processes. The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from the Business Process Simulation environment.

The paper is structured as follows. In Section 2, we review related work. In Section 3, the proposed genetic algorithm framework based on CPN is introduced. In Section 4, a case study on resource allocation case from Macau is given. In Section 5, the experimental results are collected and analyzed in detail. In section 6, we conclude the paper with future work.

2 Related Work

Although simulation experiments are widely used in evaluating business processes, to the best of authors' knowledge, no work has been reported on designing generic evolutionary framework for simulating business processes.

Simulation based analysis of business processes were reported in [2], [3]. In [9], a discrete-event simulation model was used to analyze the business process at the Center for Social Work in Slovenia for predicting the effects of the new organizational scheme, the duration of the processes, and potential bottlenecks. Escalation strategies for business processes are evaluated using simulation models in [4]. Escalation actions are used to reduce the deadline violations, or to negotiate an extended deadline with the customer. In [4], Paganos et al. proposed two strategies, which are 1) minimizing the slack time and 2) deadline prediction in order to minimize the number of escalations needed during workflow execution and to mitigate their associated costs. Simulation experiment for an insurance claim process from an Australian Company is also conducted by van der Aalst et al. In [5], they analyze various deadline escalation strategies. In their approach, escalation strategies are evaluated from three perspectives: the process perspective of using alternative path selection, the data perspective of using data degradation, and the resource perspective of using resource redeployment. In [6], four escalation strategies from [4], [5] are evaluated from temporal (workflow time) and cost (execution, resource, compensation) perspectives. In [7], a genetic algorithm is used to find near-optimal resource allocation scheme and the event-driven schedule of a Color Petri Nets.

3 Process Simulation Framework Based on CPN

In this research, we propose a generic GA modeling framework which is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and the process models. Based on the extracted information, the GA modeling framework constructs chromosomes and fitness functions required for the evolution process. The overall scheme of the generic GA modeling framework for BPS is depicted in Figure 1. First, the user imports the process model to the GA framework for analysis. Next, the GA framework uses the resource information of the model to form the members of the population for the 1st generation. Each chromosome represents a potential resource allocation scheme. For each chromosome, one round of simulation is performed to calculate the degree of fitness. Chromosome fitness is calculated based on the workflow completion time and the total cost. These chromosomes are then ordered according to the fitness and genetic operators such as selection, crossover, and mutation are applied to the chromosomes to form new members for the next generation. The overall process iterates until the change in the fitness of the best members in the population for several consecutive generations is less than a predefined threshold. The resource allocation scheme encoded in the fittest chromosome from the last generation is then returned as the best resource allocation scheme for the process model. The GA modeling framework and input process models are both implemented using Color Petri Nets and simulation of process models are conducted in CPN Tools [14].

Algorithm 1 shows the pseudo-code of our general genetic algorithm framework. For different resources types, a predefined chromosome structure is used to capture their upper and lower bounds. In the proposed model, there is no limit on the number of resource types for the simulation. The predefined chromosome structure is in the form of: [(upper bound of resource n , resource n), (upper bound of resource $n-1$, resource $n-1$), ..., (upper bound of resource 1, resource 1)].

Population size depends on the nature of the problem and chromosomes are randomly generated to cover the entire range of possible solutions. In our proposed genetic algorithm framework, we define a fitness function $f(c, g)$ for calculating the fitness of chromosome c from generation g based on the total workflow completion time and total cost of all tasks n in the workflow.

$$f(c, g) = \frac{1}{\sum_1^n time_i^{c,g}} + \frac{1}{\sum_1^n cost_i^{c,g}} \quad (1)$$

Evolution process of genetic algorithm stops when a termination condition has been reached. In our case, the simulation will stop when the change in the fitness of several generations is less than a predefined threshold. The genetic algorithm then returns the fittest chromosome from the final generation. The resulted chromosome represents the best possible resource allocation scheme for the process model.

Although we have used some default settings for the simulation, they can be easily changed by the user to meet their requirements. In our prototype, the population size is set to 10 in each generation and there is no limit on the number of resource types for the process model. The upper bound of a resource can be defined in the

chromosome structure and the selection and elimination rate is set to 90% (40% for crossover and 50% for mutation) and 10% respectively. The crossover point is randomly generated by the GA CPN model and the mutation rate is $1/L$ where L is the length of the chromosome. All these settings are stored in the *places* in the Color Petri Nets and can be altered according to the user's requirement.

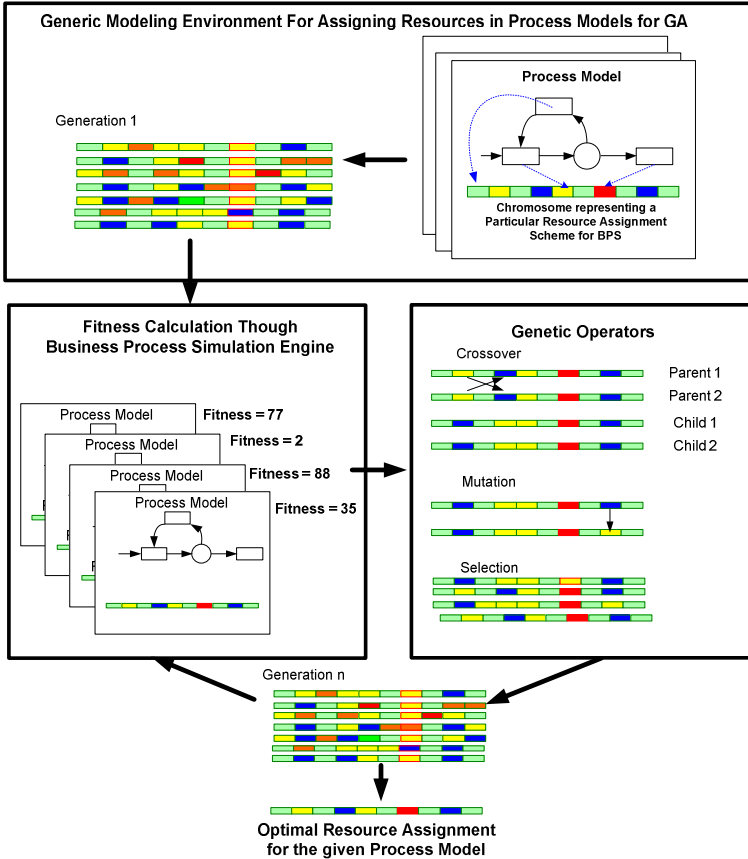


Fig. 1. The overall scheme of the generic Genetic Algorithm modeling framework for resource assignment in Business Process Simulation

The main contribution of the proposed framework is that, the process models depicted in the upper right hand corner of Figure 1 can be substituted with any Color Petri Nets process models. Since the remaining modules of Genetic Algorithm modeling framework in Figure 1 are also implemented in Color Petri Nets, these modules can be seamlessly integrated with any process model for simulation (see Figure 7 from Appendix).

To enable this integration, we define an interface based on *places* in the Petri Nets model for importing and exporting of chromosomes from the business process model

to the GA algorithm. These chromosomes are programmed as *tokens* in the Petri Nets model and stored in the places (depicted as ellipses). The high level processes “Pass in” and “GA” which are depicted as rectangles with double border lines in the highlighted area of Figure 7 are implemented as separate Petri Nets models and encompass all the modules of the proposed Generic Modeling Framework. These modules communicate with the business process model which is programmed as transitions and places on the left side of Figure 7 via the places called “Pool” and “Results”. In Figure 7, the place “Pool” which belongs to the business process model contains tokens for the simulation. These tokens are referenced during the simulation by the respective tasks in the business process model. After the simulation, the result is returned to the GA modules from the process “Measurement system” via the place “Results”. In our prototype, all transitions (rectangles with double border lines) in Figure 7 are implemented as separate Petri Nets models. Due to the limited space, only the highest level of Petri Nets model is shown in the Appendix.

```

procedure genetic algorithm
begin
    set generation g:=1;
    set i = 1;
    initialize the population P(g) = pop;
    Repeat
        While  $i \leq \text{pop}$  then Do
            run the simulation based on chromosomei;

            evaluate fitness of chromosomei;

            sort the chromosomei by fitness;
            i:=i+1;
        end
        select 40% of parents from the population P(g);
        crossover to produce offspring from these pairs;
        mutate the remaining 50% of the candidates of P(g);
        randomly generate 10% candidates for the new population;
        replace the weakest candidate of P(g) with these new
        offspring in P(g+1);
        set generation g := g+1;
        set i:=1;
    Until change in degree of fitness <= predefined threshold
    return the chromosome with best fitness result;
end

```

Algorithm 1. Pseudo-code of genetic algorithm CPN model

4 Case Study

In this section, we demonstrate the application of the proposed framework based on an archival management workflow. Archive [1] is a collection of records and documents which need to be kept and conserved as an invaluable asset. We conceptualize the workflow model of Macau Historical Archives in Figure 2. Color Petri Nets representation of the workflow model is depicted in Figure 7 of Appendix. First, records for archiving are evaluated in the “Appraisal” (T1) process. Appraisal is the process of assessing whether these records have sufficient value to warrant

acquisition by an archival institution. After the appraisal task, these records are formally accepted by the archive institute in “Receive record transfer” (T2) process. Next, in “Initial conservation” (T3) process, basic cleaning of the records is performed before they are grouped in certain order in “Arrangement” (T4) process. Then “Description” (T5) process is carried out to analyze, organize and record details of the archive based on international description standards. After descriptions are added, “Paper surrogating” (T6) process is carried out to create digital archives from the physical format. The next steps in the process involve creating backups and rebinding archives. In “Backup” (T8) process, the original files are copied into storage media so that it can be restored if the original data is deleted or damaged. “Rebinding” (T9) process is performed for repackaging as well as for associating the related meta-information with the specific archival record for permanent storage. “Microfilm surrogating” (T7) process captures and stores images of the archives in microfilm formats. Finally these records are stored at the permanent storage. Depending of the nature of the archived material, periodic maintenance tasks (i.e. preservation) are also scheduled at the end of the process.

In Macau Historical Archives, at the beginning of each year, transfer lists are sent to the Macau Historical Archives by various organizations. A transfer list contains

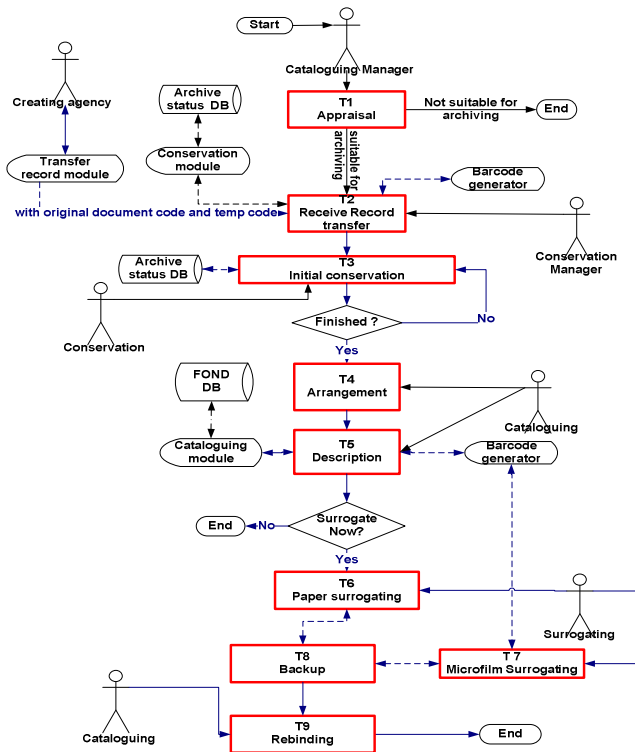


Fig. 2. Archival management workflow (Level 0)

approximately 225 records for archiving. In normal condition, approximately two transfer lists and a box of microfilm are received per year. Each box of microfilm contains approximately 2000 images. According to the recent statistics in Macau Historical Archives, approximately 14% microfilm pages are delayed and the bottleneck was located at T4, T5 and T6 since relatively high number of assigned records cannot be processed on time. We also find that delays in these tasks have a ripple effect on the whole workflow process. In the GA CPN models, each task is assigned with an “estimated average completion time” for processing job (see Table 1). These values are calculated from the recent statistical data from Macau Historical Archives. We can see that the total completion time of the workflow is approximately 1580 minutes.

Table 1. Average completion time of each task in the workflow

Task	Task description	Estimated average task cost (per record)	Estimated average completion time (per record)
T1	Appraisal	\$5.28	2 minutes
T2	Received record transfer	\$9.5	4 minutes 25 seconds
T3	Initial conservation	\$182.83	85 minutes
T4	Arrangement	\$4.3	2 minutes
T5	Description	\$3810.15	24 Hours
T6	Paper surrogating	\$37.23	19 Minutes 30 seconds
T7	Microfilm surrogating	\$12.84	6 Minutes
T8	Backup	\$17.83	8 Minutes 20 seconds
T9	Rebinding	\$18.43	10 Minutes
Total (per record)		\$4098.39	1577 Minutes 15 seconds

For better accuracy, each simulation experiments are run for 4 years period and average values are calculated for comparison. Also in this experiment, there are 14 types of resources and constraints on these resources are defined in the chromosomes. The chromosome [(2,14), (2,13), (1,12), (1,11), (1,10), (1,9), (7,8), (4,7), (6,6), (8,5), (1,4), (1,3), (1,2), (1,1)] is used to encode the resource assignment scheme which is currently used in Macau Historical Archives. For illustration, we denote this scheme as “original as-is model”.

5 Experimental Result

From the experiment result (Figure 3, 4, 5) we can see the original as-is model has relatively long average waiting time for each task. It also has the longest workflow time and largest total workflow cost among all experimental models. The resource allocation schemes found by GA from last five generations are depicted as G26, G27, G28, G29, and G30. The fittest chromosome found in 30th generation is: [(3,14), (2,13), (3,12), (2,11), (3,10), (3,9), (4,8), (13,7), (1,6), (9,5), (2,4), (1,3), (1,2), (1,1)]. In Figure 3, we can see that T4, T5, T6 and T8 have relatively high average waiting time.

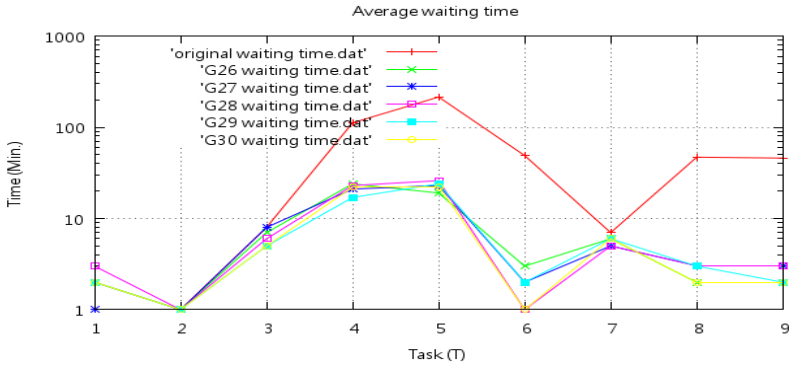


Fig. 3. Average waiting time

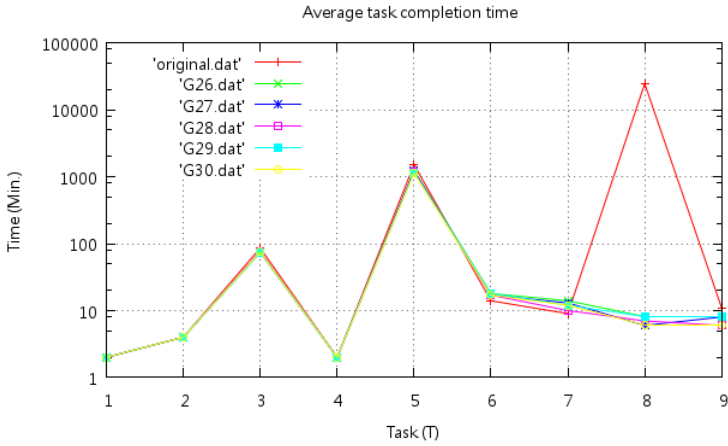


Fig. 4. Average task completion time

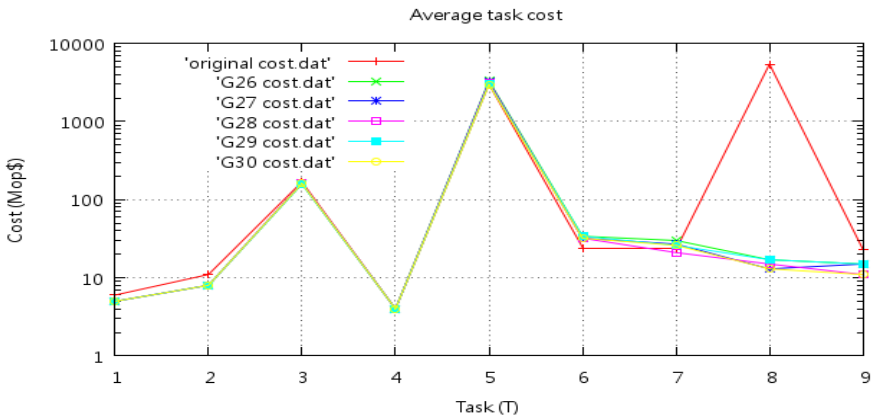


Fig. 5. Average task cost

In figure 6, experiment results show that the estimated average total workflow time and average total workflow cost gradually decrease in later generations.

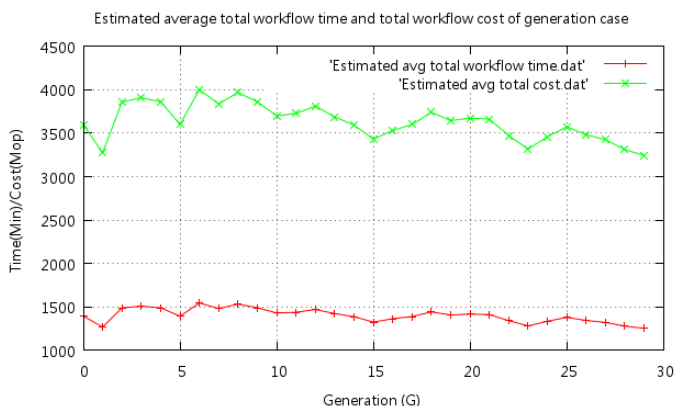


Fig. 6. Estimated average total workflow time and average total workflow cost for each generation

From the above experiments, we find that GA-based approach significantly reduces workflow time and the total cost compared to original resource allocation scheme. We can see that original as-is model has the longest workflow time among all experimental models. The performance of GA-based resource allocation in generation 30 achieves 20% reduction in total workflow time and 21% reduction in total workflow cost.

6 Conclusion

In this paper, we detail a generic evolutionary framework which can be used to simulate different kinds of business processes. The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from a given business process and can produce the best resource assignment scheme. The framework was developed in Color Petri Nets and can be used to simulate various workflows for identifying the best resource allocation scheme. The use of the framework was evaluated against a case study on Macau Archival Management Workflow.

So far, the proposed generic evolutionary framework can be used to simulate any process models defined in Color Petri Nets. As for the future work, we are planning to extend our framework for simulating workflows which are defined in other process modeling notations.

Acknowledgments. This research is funded by the University of Macau.

References

1. Pearce-Moses, R.: SAA: A Glossary of Archival and Records Terminology. The Society of American Archivists, <http://www.archivists.org/glossary/index.asp>
2. Kovacic, A., Pecek, B.: Use of Simulation in a Public Administration Process. *Simulation* 83(12), 851–861 (2007)
3. Greasley, A.: Using Business-Process Simulation within a Business-Process Reengineering Approach. *Business Process Management Journal* 9(4), 408–420 (2003)
4. Panagos, E., Rabinovich, M.: Reducing Escalation-Related Costs in WFMSs. In: *Proceedings of NATO Advanced Study Institute on Workflow Management Systems and Interoperability*, pp. 107–127. Springer, Heidelberg (1998)
5. van der Aalst, W.M.P., Rosemann, M., Dumas, M.: Deadline-based Escalation in Process-Aware Information Systems. *Decision Support Systems* 43(2), 492–511 (2007)
6. Chan, K.L., Si, Y.-W., Dumas, M.: Simulation-based Evaluation of Workflow Escalation Strategies. In: *Proceedings of the 2009 IEEE International Conference on e-Business Engineering (ICEBE 2009)*, pp. 75–82. IEEE Press (2009)
7. Saitou, K., Malpathak, S., Qvam, H.: Robust Design of FMSs using CPN and Genetic Algorithm. *Journal of Intelligent Manufacturing* (13), 339–351 (2002)
8. An, L., Jeng, J.-J.: On developing system dynamics model for business process simulation. In: *Proceedings of the 37th Conference on Winter Simulation (WSC)*, pp. 2068–2077 (2005)
9. Lee, Y.M.: Simulating Availability Outlook for E-Commerce Business for Personal Computer Sales. In: *Proceeding of the 36th Conference on Winter Simulation (WSC)*, pp. 1201–1204 (2004)
10. Book, M., Gruhn, V., Hülder, M., Köhler, A., Kriegel, A.: Cost and Response Time Simulation for Web-based Applications on Mobile Channels. In: *Proceedings of the 5th International Conference on Quality Software*, pp. 83–90 (2005)
11. Holland, J.H.: *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor (1975)
12. Lee, H., Kim, S.-S.: Integration of Process Planning and Scheduling Using Simulation Based Genetic Algorithms. *The International Journal of Advanced Manufacturing Technology* 18(8), 586–590 (2001)
13. Hegazy, T., Kassab, M.: Resource Optimization Using Combined Simulation and Genetic Algorithms. *Journal of Construction Engineering and Management* 129(6), 698–705 (2003)
14. CPN Group, University of Aarhus, Denmark. CPN Tools Home Page, <http://wiki.daimi.au.dk/cpntools/>

Test-Activity Flow Control for SOA Execution Test

Youngkon Lee

e-Business Department, Korea Polytechnic University,
2121 Jeongwangdong, Siheung City, Korea
yklee777@kpu.ac.kr

Abstract. Electronic collaboration over the Internet between business partners appear to be converging toward well-established types of message exchange patterns that involve both user-defined standards and infrastructure standards. At the same time, the notion of event is increasingly promoted for asynchronous communication and coordination in SOA systems. In collaboration between partners or between components is achieved by the means of choreographed exchanges of discrete units of data - messages or events - over an Internet-based protocol. This paper presents an event-centric test case scripting method and execution model for such systems.

Keywords: SOA, event-driven, process invocation.

1 Introduction

While current Web Service technologies show much progress, current services are mainly limited to atomic services. Thus, they are not adequate to handle the autonomous and complex services in realistic settings. In dealing with this problem, some research works have developed languages to compose the individual Web Services into transactions or workflows. Web Services Flow Language (WSFL) [1] was designed for service compositions in the form of a workflow, and XLANG [2] for the behavior of a single Web Service. However, these works are not sufficient for providing the adaptive web Services generated from a particular context.

Electronic collaborations over the Internet between business partners (e-Business / e-Government) appear to be converging toward well-established types of message exchange patterns that involve both user-defined standards and infrastructure standards. At the same time, the notion of event is increasingly promoted for asynchronous communication and coordination in Event-Driven Architectures (EDA) that are considered as either complementary to or part of SOA systems. In both cases collaboration between partners or between components is achieved by the means of choreographed exchanges of discrete units of data - messages or events - over an Internet-based protocol. Such systems require an event-centric test case scripting markup and execution model.

In e-Business transactions as in EDAs, partners or components must agree on the use of a combination of standards in order to interoperate with each other. Typically, these standards can be classified into three layers:

- Messaging infrastructure standards, ranging from transport level to higher-level messaging protocols and quality of service (QoS) including reliability and security, such as those defined as SOAP extensions, or REST (Representational State Transfer).
- Multi-message exchange standards as manifested in business processes and choreographies.
- Business document standards may be business content structure and semantics, taxonomies in use, code lists, semantic rules, or the XML schema modeling style. They are industry-specific (e.g. RosettaNet PIP schemas, AIAG Inventory Visibility and Interoperability schemas), horizontal document standards, or regional guidelines.

There have been conformance and interoperability test suites and testing tools for above layer individually. But the testing of integrations of standards has been ad-hoc, or limited mostly to standards in the messaging infrastructure.

Although the need for testing some form of integration of standards has been well recognized for infrastructure standards, there has been little support for testing integrations that extend to the use of standards specific to a business - e.g. for documents or choreographies. Such integrations can be construed as user-defined profiles. For example, the level of QoS required for a business transaction may depend on the nature of business data being exchanged, or on some property defined by the related business process.

Testing and monitoring these infrastructure layers and their integration also requires that test cases access a combination of contracts - agreements, policies or business transaction patterns - represented by meta-level documents.

This compliance objective goes beyond quality assurance for the messaging function: it requires the monitoring of live transactions in production environments, as well as verifying conformance of business endpoints in operation conditions. This calls for a flexible test execution model that can accommodate performance constraints as well as different timeliness constraints - e.g. where tests are either deferred over log data, or executed on live exchanges in a monitoring mode.

Consequently, the execution model of such test cases or monitoring cases, must accommodate dynamic conditions requiring real-time or near real-time error detection or measurement, allowing to correct and report on business exchanges as they proceed.

The output of a monitoring script also must provide more information than a report of the type pass / fail. Different ways of "passing" or "failing" must be reported on, as well as identifying the types of business transactions. The output must be easy to format and feed to another decision engine, such as a rule engine that will process this input in real-time. For example, a rule may decide to generate an alert if a business transaction lasts too long, depending on the nature of the transaction and on the SLA associated to these business partners.

This paper defines a testing and monitoring model, as well as a test script markup, so that test cases or monitoring cases can be fully automated, and portable across test environments. In section 2, we summarize the related works regarding the web service flow language for testing. Section 3 presents the concept of Event-Centric

Test Case Script (EVEC), section 4 describes the implementation of EVEC, and we conclude in section 5.

2 Related Works

In fact, the automatic or semi-automatic management of service flows over the Web has not been achieved yet. In the Web Services model that is quite different from traditional one, there are a large number of similar or equivalent services which user can freely select and use for their application. Since the service is developed and deployed by the third party, the quality of service is not guaranteed. The services may not be adequate as per service requestor's requirements and kept evolving, without notification to service requestors, according to the provider's requirements and computing environment.

Thus, it is important to provide adaptability to evolving services as well as diverse context of services. Kammer et al. [3] suggested workflow to be dynamic, which allows changes with minimal impact to the ongoing execution of underlying workflow, as well as be reflexive, which provides knowledge about a workflow's applicability to the context and the effectiveness of its deployment evaluated over time. Understanding constraints and context associated with services may affect the quality of service. From this perspective, optimization may occur through the evaluation and refinement of a previous service flow.

Automatic composition of services is challenging, because it is difficult to capture semantics and context of services and measure the quality of services. One exemplary effort that aims for this function is DAML-based Web Service Ontology (DAML-S) [4], which describes the properties and capabilities of Web services.

Workflow technology has been around since a decade ago and has been successful in automating many complex business processes. A significant amount of work has been done in this field, which deals with different aspects of workflow technology process modeling, dynamic workflows, and distributed workflows. Process modeling languages such as IDEF, PIF, PSL or CIMOSA [5] and frame based models of services were used to design process typing, resource dependencies, ports, task decomposition and exception.

Current research on web services paves way for web service based workflows, which has obvious advantages pertaining to scalability, heterogeneity, reuse and maintenance of services. Major issues in such inter-organizational service based workflows are service discovery, service contracts and service composition. Web Services Flow Language (WSFL) was proposed to describe compositions of services in the form of a workflow, which describes the order of service invocation. Service composition aids such as BizTalk [6] were proposed to overcome the limitations of traditional workflow tools which manually specify the composition of programs to perform some tasks. Other industrial initiatives such as BPEL4WS [7], and XLANG concentrates on service representation issues to tackle the problems of service contracts, compositions and agreements. Current efforts are to automate the complete

process of service discovery, composition and binding, using the machine understandable languages. Some other recent advances are WS-Transaction and WS-Coordination which define protocols for executing Transactions among web services. There is a research for modeling QoS of workflows [8], and defining a QoS based middleware for services associated with the underlying workflow [9], but it doesn't take into account QoS factors related to Internet based services. Some researchers describe QoS issues related to web services from the provider's perspective [10]. We believe that the current research has not delved into QoS issues related to Web Service based workflows, and many critical issues related to the availability, reliability, performance and security of Web Services need to be handled. Our approach tactfully utilizes and monitors these QoS parameters to provide a consistent service interface to other applications in the workflow through adaptive QoS based selection, binding and execution of Web Services.

3 Event-Centric Test Case Script

Event-centric test case script (EVEC) is designed so that the same scripts can be used either in live monitoring mode, or in analysis of past events from a log (referred to as deferred mode in this paper called hereafter the "deferred mode") or yet in mixed situation. Testing and Monitoring of Business Processes as well as more generally of systems the behaviour of which can be traced by events, fall in the following three categories:

- Compliance with specifications. Such specifications may be of a business transaction, business process definition, documents exchanged, or of infrastructure behaviour (e.g. messaging protocol). Enabling the automatic generation of EVEC scripts from such specifications when these specifications are formal – e.g. process definition, choreographies, document schema or rules – is part of the requirements although the methodology to achieve this is out of scope of this document. Some test assertion design and best practices, such as those in Test Assertions Guidelines [11] may be used for deriving scripts from such representations even when automatic generation is not possible.

- Compliance with agreements. Such agreements may be business agreements such as SLAs, or regulatory compliance rules. They may also be infrastructure configuration agreements (e.g. ebXML CPA, WS-Policy). This category of application includes SLA monitoring, business metrics and aspects of business activity monitoring (BAM) that are closest to operations, e.g. for regulatory compliance.

- Business Operation intelligence. Such monitoring is not directly related to compliance, but primarily intended for generating reports and various analytics of business activities. This includes analyzing the logs of processes and business transactions for reports and BI. This category of application includes BAM (business activity monitoring). In its dynamic aspect, this monitoring is addresses the need for visibility in business processes and service-oriented systems, which include problem detection/anticipation, diagnostics and alarm generation. Each one of the above categories may be considered both in a real-time context (e.g. generation of alarms

and notifications during operation) and a deferred, off-line analysis context (periodic generation of reports or metrics with no direct, automatic feedback loop to operations). In both cases, the same input – in form of events – is assumed.

From the viewpoint of script execution semantics, "live" and "deferred" are not distinguished: the same script is executable on input that is either live or logged. To ensure flexibility for handling various monitoring contexts and situations, mixing of both execution modes must be supported:

- A script may start executing "deferred mode" with its events already partially logged, and then catch-up with the on-going logging of events and continue "live".
- Conversely, a script may start live, and if its execution engine is interrupted for some reason, may resume its analysis of events that have already been logged while the test engine was stopped, in deferred mode. Then it may catch-up with events and eventually go live again. When events are consumed in a publish-subscribe mode, a simple queuing mechanism is sufficient to provide the above flexibility. However, EVEC must be able to correlate with past events.

4 Implementation of EVEC

The EVEC script language is designed for testing and monitoring processes or business transactions of various kinds, and more particularly for analyzing and validating event patterns that are generated by these processes. To this extent, EVEC may be described as an event-processing language. The content and structure of these events may be quite diverse, but a common XML wrapper is assumed. The top-level constructs are the script package and the scriptlet:

- The script package, or "script": This is the main unit of execution. The script package contains an "execution context" (<execution-context> element) that defines various global constructs and bindings, e.g. for declaring event boards. The script package also contains one or more "scriptlets". The execution context in a script package defines which scriptlet to start execution with - or main scriptlet. In case other scriptlets are defined, the main scriptlet is expected to invoke these directly or indirectly.
- The scriptlet: A scriptlet defines a flow (or thread) of execution for a sequence of atomic operations. Scriptlets can execute either concurrently or not (see detailed meaning in the next section), and can be started in a blocking or non-blocking way.

EVEC is designed so that it leverages existing XML script languages for special features such as logical conditions and event selection. The default language for all logical expressions over XML content is XPath, along with its existing function libraries (e.g. advanced set of functions for time and duration management).

The concept of concurrency in EVEC is entirely dependent on the notion of "virtual present time" (VPtime). When a scriptlet starts to execute, it is assigned a VP-time which will condition its event consumption and timestamp its event production. The default VP-time assignments are:

- The first scriptlet of a script package is assigned the initial VP-time of this script, the default of which is in turn the actual present time (AP-time).
- The VP-time of a scriptlet S2 started by a scriptlet S1, is the value of VP-time in S1 when [start S2] is executed. These default values can be overridden by the <start> operation, which allows to set the VP-time of the started scriptlet (see the start/@vptset attribute in section 4). Inside a scriptlet, the VP-time may be progressed by two operations:
 - <wait> : will add some predefined duration to the VP-time, or wait until some date, or yet until some other scriptlets complete.
 - <catch> : when waiting - virtually or really - for some event to occur, will advance the VP-time to the occurring date of events being caught.

Event catching in a scriptlet is only considering (by default) events occurring at or after the current VP-time. Besides <wait> and <catch>, the execution duration of other EVEC operations is considered as negligible, as far as the VP-time is concerned: in other words, these operations do not affect the VP-time. The VP-time window of a scriptlet execution is defined as the [starting VP-time, ending VP-time] time interval of the execution. Intuitively, concurrent execution is achieved when the VP-time windows of two scriptlets overlap. Concurrency is entirely determined by the start VP-time assigned to these scriptlets. When a scriptlet S1 starts a scriptlet S2, it can do so either in a blocking or non-blocking way:

- **Blocking invocation:** intuitively, the invoking scriptlet (S1) will wait until the invoked scriptlet (S2) terminates. The next statement in the invoking scriptlet S1 (after <start>), will execute at a VP-time that is same as the VP-time set at the end of the invoked scriptlet (S2). In other words, the VPtimes of S1 and S2 are “synchronized” after S2 completes (see start/@vptsync="true" in Section 4). More generally, to accommodate the case where a starting time is set at a date/time anterior to the

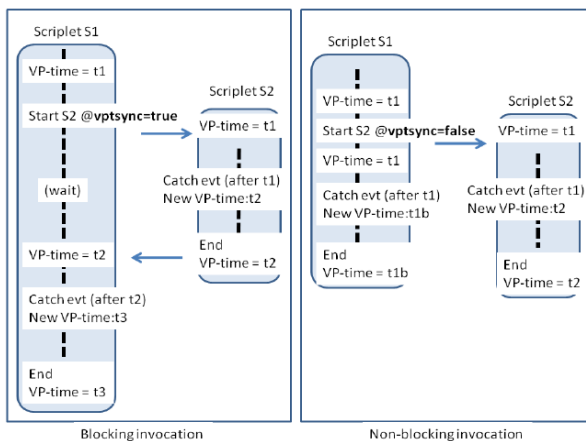


Fig. 1. Blocking and non-blocking modes of scriptlet invocation

invoking time (@vptset="a past date/time"), the VP-time of the next statement in S1 is either the last VP-time value of S2 or the last VP-time value in S1 (just before invoking S2), whichever occurs the latest.

- Non-blocking invocation: intuitively, the invoked scriplet (S2) will not cause the invoking scriplet (S1) to wait. In other words, the VP-times of S1 and S2 are not "synchronized" after S2 completes. (see start/@vptsync="false" in Section 4). The next statement in the invoking scriplet S1 will execute at a VP-time that is same as the VP-time value just before executing the <start> statement, this regardless of the value of start/@vptset. Non-blocking invocations should not be seen as only useful for actual concurrent (or multi-threaded) processing. In many cases, it makes scripting easier and more intuitive, even when execution is entirely deferred on past (logged) events that could otherwise be processed serially in a single-threaded way. Various cases of blocking and non-blocking invocations are illustrated below. The following figure illustrates both modes of scriplet invocations, and how the VP-time is affected – or not - in the invoking scriplet.

When a scriplet S1 does a blocking invocation of a scriplet S2, the VP-time of the invoked scriplet S2 is initiated at the current VP-time of the invoking scriplet S1 (unless a different VPtime value is given using start/@vptset as illustrated in the next figures). The scriplet S1 is then "blocked" until the VP-time at the end of S2 is known and assigned as current VP-time in S1.

In the non-blocking invocation (see Fig. 1.), S1 will ignore the ending time of S2. A single-threaded execution may still execute S2 first before executing the next statement in S1. The execution semantics would still allow "concurrent" catching (in terms of virtual present) of same events by S1 and S2, as both will select events starting from VP-time t1. In the above figure, S1 is catching an event at time t1b

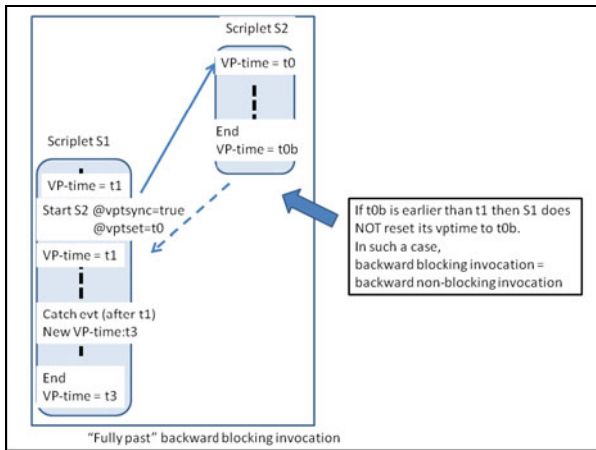


Fig. 2. Fully past backward blocking invocation

while S2 is catching an event at time t_2 . Depending on their respective selection expressions, these catches could capture either the same or different event, causing the new VPtime in S1 (t_{1b}) to be either prior or after the new VP-time in S2 (t_2).

In a single-threaded execution of the non-blocking case that started “live” (VP-time = present time), S2 could be executed first live, then the remaining part of S1 can be executed “deferred” on the log of events, starting from time t_1 now in the past. Clearly, more fine-grain serialization of S1 and S2 executions would be required if these two scriptlets communicate with each other, e.g. if S1 is consuming an event posted by S2 or vice-versa.

5 Conclusion

In this paper, we present an event-centric test case scripting method and execution model, EVEC, for such systems. EVEC enables testing and monitoring of applications and business processes, the behavior of which can be traced and monitored via events. The notion of event in EVEC is broad, thus, it could cover all of the type of SOA business processes. Further study is required to define and classify the detailed test case metadata or artifacts that would complement EVEC in test environments.

References

1. Leymann, F.: Web services flow language, TR WSFL 1.0. IBM Software Group (May 2001)
2. Thatte, S.: XLANG Web Services for Business Process Design (2001), <http://www.gotdotnet.com/team/xmlwsspecs/xlang-c/default.htm>
3. Kammer, P., Bolcer, G.A., Taylor, R.N., Bergman, M.: Techniques for Supporting Dynamic and Adaptive Workflow. *Journal of Computer Supported Cooperative Work (CSCW)* 9, 269–292
4. DAML-S Specifications, <http://www.daml.org/services/>
5. Kosanke, K.: CIMOSA - Open System Architecture for CIM. In: ESPRIT Consortium AMICE. Springer, Heidelberg (1993)
6. Biztalk, <http://www.microsoft.com/biztalk/>
7. Business Process Execution Language for Web Services, Version 1.0 (July 2002), <http://www-106.ibm.com/developerworks/webservices/library/ws-bpel/>
8. Cardoso, J., Sheth, A., Miller, J.: Workflow Quality Of Service (2002)
9. Sheth, A., Cardoso, J., Miller, J., Koch, K.: QoS for Service-oriented Middleware. In: *Proceedings of the Conference on Systemics, Cybernetics and Informatics, Web Services and Grid Computing, Orlando, FL (July 2002)*
10. Mani, A., Nagarajan, A.: Understanding quality of service for Web services, <http://herzberg.ca.sandia.gov/jess/>
11. OASIS Test Assertions Guidelines (TAG) TC, “Test Assertions Guidelines”, <http://www.oasis-open.org/committees/tag/>

Business-Rating Based SLA Model for SOA Quality

Youngkon Lee

e-Business Department, Korea Polytechnic University,
2121 Jeongwangdong, Siheung City, Korea
yklee777@kpu.ac.kr

Abstract. With Web services starting to be deployed within organizations and being offered as paid services across organizational boundaries, quality of service (QoS) has become one of the key issues to be addressed by providers and clients. While methods to describe and advertise QoS properties have been developed, the main outstanding issue remains how to implement a service that lives up to promised QoS properties. This paper provides the service level agreement (SLA) parameters for QoS management applied to Web services and raises a set of research issues that originate in the virtualization aspect of services and are specific to QoS management in a services environment – beyond what is addressed so far by work in the areas of distributed systems and performance management.

Keywords: SOA, business context, SLA parameter.

1 Introduction

Whether offered within an organization or as a part of a paid service across organizational boundaries, quality-of-service (QoS) aspects of services are important in a service-oriented computing environment. Dealing with QoS is a sign of a technology going beyond its stage of initial experimentation to a production deployment and many recent activities related to QoS of Web services indicate that this is becoming an increasingly relevant topic.

Efforts in the past years mainly focused on describing, advertising and signing up to Web and Grid services at defined QoS levels. This includes HP's Web Services Management Framework (WSMF) [1], IBM's Web Service Level Agreement (WSLA) language [2][3], the Web Services Offer Language (WSOL) [4] as well as approaches based on WS-Policy [5]. These efforts enable us to describe quality metrics of services, such as response time, and the associated service level objectives flexibly and in a way that is meaningful for the business needs of a service client.

However, one of the challenging issues is to associate or derive a system configuration that delivers the QoS of a described Web service using the above mentioned approaches. In many cases, this is non-trivial. Sometimes we can rely on experience with tested, dedicated system configurations to decide, for example, the size of a cluster for a particular workload guaranteeing a particular response time for a given percentile of requests. In addition, managing a service at different QoS levels on the same infrastructure is not easy.

While managing QoS in distributed systems is not a novel problem, a number of additional issues arise in the context of a service-oriented computing environment. Those issues arise from the specific properties of Web services. For example, cross-organizational Web services may be accessed through the public Internet and client side QoS metrics have to include network properties in addition to properties of the service-implementing application itself. In addition, composite and recursively aggregated services – and the ability to aggregate is seen as a key benefit of Web services – gain new QoS properties that are not always easily derived from their parts.

The objective of paper is to analyze the main QoS factors of Service-oriented Architecture (SOA) in a business context and to provide the service level agreement (SLA) parameters that affect critically the business performance.

According to OASIS Reference Model for Service Oriented Architecture [SOA-RM] [6], the Service Oriented Architecture (SOA) is a paradigm for organizing and utilizing distributed capabilities that may be under control of different ownership domains. The service within SOA is a mechanism to enable access to one or more capabilities, where the access is provided using a prescribed interface and is exercised consistent with constraints and policies as specified by the service description. This specification further defines the business service level agreement (bSLA) between the service requester and the service provider for the service which is defined in SOA-RM, within the end-to-end resource planning (EERP) technology [7]. The applications of EERP are any kind of business services, and they are not limited to Web Services only. This applies the well-known technique for service discovery and optimization in a novel way to improve business results. It models the business process and the range of potential services, and then guides the selection and deployment of services based on the end-to-end business value. Modeling the business service-level agreements to manage and evaluate services and establishing agreements about the business service is essential to long-term value chain improvement. The bSLA is different from the SLA in the software/IT world. The bSLA is the contact between the service requester and the service provider, and the SLA is the contract between the service provider and the network/system provider. The SLA is network/system oriented agreement that deals with network performance and system availability. The bSLA is a business oriented agreement that deals with price, time to deliver, and the quality/rating of the service.

In section 2, we summarize the related works about web service selection based on the QoS metrics. Section 3 presents the service process model which enables service performance optimization in the business respect and section 4 details the bSLA model including parties, parameters and obligations, and we conclude in section 5.

2 Related Works

The first step to manage Web service's quality is to define it. While this is important for Web services as well as in traditional distributed systems, explicit definition is particularly important in an environment transcending organizational boundaries. Quality is expressed referring to observable parameters relating to a non-functional

property, for example, the response time of a request. A level of quality is agreed upon as a constraint over those parameters, potentially dependent on a precondition. Hence, the party offering a Web service, in agreement with its customers and users, will define the QoS parameters and the particular instances of the service to which these parameters relate. In the case of a Web service, a parameter such as response time can relate to an individual invocation of an operation or a class of operations, all having the same (individual) quality properties of having an aggregate property, e.g., the average response time of this class of operations or another stochastic metric.

A further step in managing Web services QoS is the definition of the semantics of the QoS parameters. A Web service and its subscribers and users must understand what it is meant. It is important what is measured where. For performance-oriented metrics this can be at different points, as Fig. 1 illustrates.

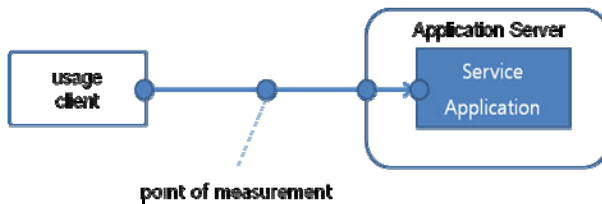


Fig. 1. Points of measurements defining semantics of metrics

The definition of QoS parameters corresponds to the establishment of ontology between a service provider and its clients. Ontology can be established in two approaches. (1) It can be a definition of terms and, potentially, or the semantics of the relationships between them, as facilitated by DAML and OIL [8]. This approach results in a fixed set of well understood terms – in our case the QoS parameters. (2) Another approach uses constructive ontology. Based on a set of well-know defined terms (as in 1) and a set of well-know composition operators, new terms' (QoS) parameters can be defined by composing new parameters out of existing ones using the operators.

Having established common understanding of quality of service parameters and the associated guarantees given by the provider, it also has to be established to which relationships between a client and a server a QoS guarantee applies. A service may provide the same quality to all requesting clients, to each client individually or to a defined set of clients that a provider organization and an organization requiring a QoS level for multiple clients agree upon in a contract, which is also called an SLA.

Clients will refer to the contract when requesting the service according to a particular quality level. The different scoping approaches of QoS guarantees require different means of establishing a particular quality level for a client: If a QoS level is associated with a service, a client searches for a suitable service in a directory, e.g., UDDI and retrieves its quality definition, e.g., stated as a WS-Policy expression. In the case of an individual client or a contract, a negotiation mechanism, which can be very simple, must be provided. Once the contract is established, the provider

organization must provision a service-implementing system such that it behaves as it has been agreed upon. This involves deriving the amount of resources needed and the runtime management of resources.

However, this is not simple. While we have developed – improvable – approaches to the issues raised above, the issue of provisioning and runtime managing a Web service-implementing system is not equally well understood yet. In the next section, we discuss what distributed systems and performance management approaches can provide.

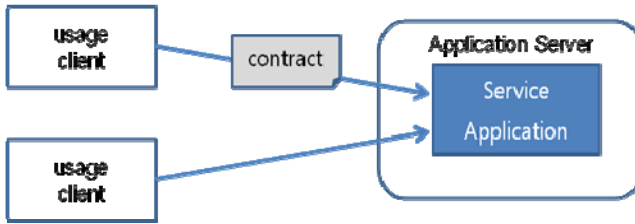


Fig. 2. Contracts defining the scope of quality guarantees

A number of performance management technologies, such as workload managers and network dispatchers, have been developed to control response times of individual systems and clusters and various availability management approaches. However, it is not straight-forward to configure, for example, workload managers to satisfy response time goals for a set of different scopes of quality – for Web services as well as for any distributed system. In this section, we outline some typical approaches of how today’s QoS parameters are managed in distributed systems.

3 Service Process Model

This section describes the service process model conceptually. Fig. 3 shows the conceptual model, and of messages flows with brief descriptions. We also include timeline and sequence diagrams Fig. 4 to show how an implementation would use service messages and build a continuous business process improvement loop. In Figure 3, the business quality of Service is abbreviated as bQoS, business rating is abbreviated as Rating, and business service level agreement is abbreviated as bSLA.

The service requester is the client system who tries to find an optimal solution provided by service optimization portal (SOP). Service providers provide business services. Each service provider may offer the same service but with different bQoS and Ratings. Services may be running on different platforms with different implementations, but they all support message exchanges of bQoS, Rating, and bSLA information in the XML formats.

The SOP accepts the request from the Service requester, performs bQoS and rating queries, calculates optimal solution(s), and then returns the result to the service requester. The Rating Provider is a party unaffiliated with either the requester

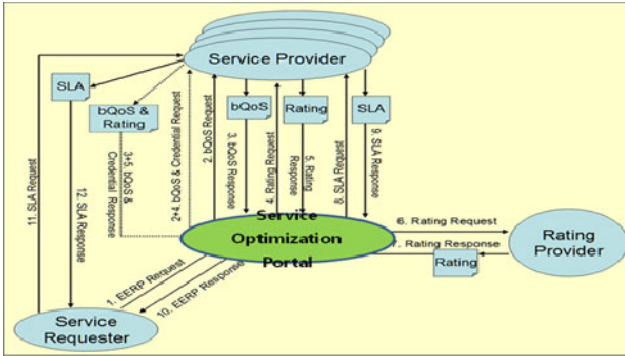


Fig. 3. Service Process Model

or the target of the rating request, such as a third party rating organization, given a reference to a particular business service and provider, issues either a number or a classification description.

There can be another way to implement the service optimization without the SOP. For example, there can be a case for some services providers and service consumers using SOA Registry-Repository to find each other, to exchange business quality of services information among them, and to begin negotiations for establishing Service Level Agreements (SLAs). The results of messages 2 through 9 in Figure 4 are used to calculate the optimal deployment for a given set of services requests. A list of alternatives might be returned in message 10. Each step in the process would have a service provider determined for each service and for each alternative. Messages 11 and 12 are exchanged between the service requester and the selected service providers to define the BSLA.

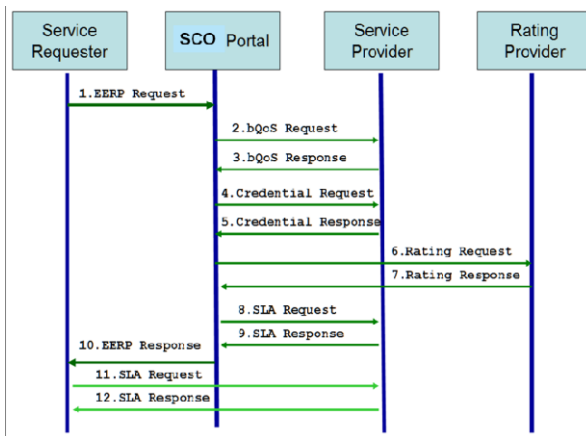


Fig. 4. Service message sequence without optional messages

4 bSLA Model

The bSLA model is for representing SLA parameters in the respect of business value. The BSLA is the root element for EERP- Business Service-level agreement (bSLA). The bSLA is a formal contract between a service provider and a client guaranteeing quantifiable business quality of service (bQoS) at defined levels. It can have one or more of the following elements:

```
<sla:BSLA xmlns:sla="..." xmlns:bqos="..." ...>
  <sla:SLAParties ...>sla:SLAPartiesType</sla:SLAParties>
  <sla:SLAParameters ...>sla:SLAParametersType</sla:SLAParameters>
  <sla:SLAObligations ...>sla:SLAObligationsType</sla:SLAObligations> ?
  <sla:SLATerms ...>sla:SLATermsType</sla:SLATerms> ?
</sla:BSLA>
```

Fig. 1. XML schema for bSLA

The following describes the attributes and elements listed in the schema outlined above:

- /sla:BSLA is a root element of Business Service-level agreement (bSLA) for EERP.
- SLAParties is a required element in bSLA that defines parties invoked in this bSLA for the service. SLAParties element has both the service provider and services requester elements.
- /sla:BSLA/sla:SLAParties/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the SLAParties element in the future. Unrecognized attributes may cause a fault or be silently ignored.
- /sla:BSLA/sla:SLAParameters is defined monitoring of bQoS metrics, including service profile uri, operations and other optional elements. It is a required element that uses sla:SLAParametersType.
- /sla:BSLA/sla:SLAParameters/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the SLAParameters element in the future.
- /sla:BSLA/sla:SLAObligations is agreed bSLA obligations aspect of the service, including obligations, action guarantees. It is a optional element that uses sla:SLAObligationsType.
- /sla:BSLA/sla:SLAObligations/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the SLA Obligations element in the future.
- /sla:BSLA/sla:SLATerms is agreed bSLA terms aspect of the service, including bSLA term elements.
- /sla:BSLA/sla:SLATerms/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the SLATerms element in the future.
- /sla:BSLA/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the root BSLA element in the future.

- /sla:BSLA/sla:BSLAExtension is an optional element that keeps different (extensible) elements to be specified in the future.
- /sla:BSLA/sla:BSLAExtension/{any} is an extensibility mechanism to allow different (extensible) elements to be specified in the future.

The SLAParties describes the list of parties invoked in the bSLA for the service. There should be one SLAParties element present in the bSLA of service. The following describes the attributes and elements listed in the schema outlined above:

- /sla:SLAParties, bSLA Parties aspect of the service, is for parties invoked in the bSLA for the service, including both service provider and service requester elements.
- /sla:SLAParties/sla:ServiceProvider represents the provider for parties. It is a required element for bSLA Parties.
- /sla:SLAParties/sla:ServiceProvider/sla:ServiceUri is a required element for Service Provider.
- /sla:SLAParties/sla:ServiceProvider/sla:ServiceProviderName is the name of the service provider. It is also a required element for Service Provider.
- /sla:SLAParties/sla:ServiceProvider/sla:ServiceProviderName/@languageID is an optional attribute in the ServiceProviderName element, using xsd:language type.

```

<sla:SLAParties xmlns:sla="..." ...>
  <sla:ServiceProvider ...>sla:ServiceProviderType
  <sla:ServiceUri ...>sla:SlaUriType</sla:ServiceUri>
  <sla:ServiceProviderName
  languageID="...">sla:ServiceProviderNameType</sla:ServiceProviderName>
  </sla:ServiceProvider>
  <sla:ServiceRequester ... >sla:ServiceRequesterType
  <sla:ServiceRequesterUri ... >sla:SlaUriType</sla:ServiceRequesterUri>
  <sla:ServiceRequesterName
  languageID="...">sla:ServiceRequesterNameType</sla:ServiceRequesterName>
  </sla:ServiceRequester>
  ...
</sla:SLAParties>

```

Fig. 6. XML schema for bSLA parties

- /sla:SLAParties/sla:ServiceProvider/@{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to ServiceProvider element in the future. /sla:SLAParties/sla:ServiceRequester represents requester for the service, including requester's name and the URI that represents the requester. It is a required element for bSLA Parties.
- /sla:SLAParties/sla:ServiceRequester/sla:ServiceRequesterUri represents the requester's identifier in URI format for the service requester. It is a required element for Service Requester.
- /sla:SLAParties/sla:ServiceRequester/sla:ServiceRequesterName is requester's name for the service requester. It is a required element for Service Requester.
- /sla:SLAParties/sla:ServiceRequester/sla:ServiceRequesterName/@languageID is an optional attribute in the ServiceRequesterName element.

- /sla:SLAParties/sla:ServiceRequester/{any} is an extensibility mechanism to allow additional attributes, based on schemas, to be added to the serviceRequester element in the future.
- /sla:SLAParties/{any} is an extensibility mechanism to allow different (extensible) elements to be specified in the future.

4 Conclusion

In this paper, we proposed a new concept of bSLA, whose items are proper to evaluate SOA service performance in the respect of service business benefit. The bSLA includes the service actor information, bSLA parties and the quality information, bSLA parameters, and bSLA obligations of the service parties. We also devised a service optimization portal which provides the best service composition by evaluating the value of bQoS of each service chain. Further study is required to define and classify the quality factor group for business case by case.

References

1. Catania, N., Kumar, P., Murray, B., Pourhedari, H., Vambenepe, W., Wurster, K.: Web Services Management Framework, Version 2.0, Hewlett-Packard (July 16, 2003), <http://devresource.hp.com/drc/specifications/wsmf/WSMF-WSM.jsp>
2. Ludwig, H., Keller, A., Dan, A., King, R., Franck, R.: A service level agreement language for dynamic electronic services. *Electronic Commerce Research* 3, 43–59 (2003)
3. Ludwig, H., Keller, A., Dan, A., King, R., Franck, R.: Web Service Level Agreement (WSLA) Language Specification, Version 1.0, IBM Corporation (January 28, 2003), <http://www.research.ibm.com/wsla/WSLASpecV1-20030128.pdf>
4. Tasic, V., Pagurek, B., Patel, K.: WSOL – A Language for the Formal Specification of Classes of Service for Web Services. In: *Proc. of ICWS 2003 (The 2003 International Conference on Web Services)*, June 23-26, 2003, pp. 375–381. CSREA Press, Las Vegas (2003)
5. Box, D., Curbera, F., Hondo, M., Kale, C., Langworthy, D., Nadalin, A., Nagaratnam, N., Nottingham, M., von Riegen, C., Shewchuk, J.: *Web Services Policy Framework (WSPolicy)* (May 28, 2003), <http://www.ibm.com/developer-works/library/ws-policy>
6. Mackenzie, C.M., et al.: *Reference Model for Service Oriented Architecture 1.0*. OASIS Committee Draft (August 2006)
7. Cox, W., et al.: *SOA-EERP Business Quality of Service Version*. OASIS Committee Draft (November 2010)
8. Connolly, D., van Harmelen, F., Horrocks, I., McGuinness, D.L., Patel-Schneider, P.F., Stein, L.A.: *DAML+OIL (March 2001) Reference Description*, W3C (December 18, 2001), <http://www.w3.org/TR/daml+oil-reference>

Semi-automatic Test Assertion Transformation Scheme for SOA

Youngkon Lee

e-Business Department, Korea Polytechnic University,
2121 Jeongwangdong, Siheung City, Korea
yk.lee777@kpu.ac.kr

Abstract. This paper presents a design method for business-centric SOA test framework. The reference architecture of SOA system is usually layered: business process layer, service layer, and computing resource layer. In the architecture, there are so many subsystems affecting the system's performance, which relates with each other. As a result, in respect of overall performance, it is meaningless to measure each subsystem's performance separately. In SOA system, the performance of the business process layer with which users keep in contact usually depends on the summation of the performance of the other lower layers. Therefore, for testing SOA system, test cases describing business process activities should be prepared. We devised a business-centric SOA test assertion model which enables to semi-automatic transform test assertions into test cases by the concept of prescription level and normalized prerequisite definition. The model also minimizes the semantic distortion in the transformation process.

Keywords: SOA, business process, test assertions, test cases.

1 Introduction

Service Oriented Architecture (SOA) is generally defined as a business-centric IT architectural approach that supports integrating businesses as linked, repeatable business tasks, or services. SOA enables to solve integration complexity problem and facilitates broad-scale interoperability and unlimited collaboration across the enterprise. It also provides flexibility and agility to address changing business requirements in lower cost and lesser time to market via reuse.

SOA has a lot of promises of interoperability, however, at the cost of: lack of enterprise scale QoS, complex standards which are still forming, lack of tools and framework to support standards, and perform penalty. Recently, as SOA has been widely adopted in business system framework, performance issues in SOA are raised continuously from users and developers.

SOA system is generally composed of various subsystems, each of which relates intimately with others. Therefore, if performances are issued, it is very difficult to find out the reason clearly. For example, if a business process in SOA system has longer response time than before, there could be various reasons: cache overflow in a

business processor, wrapping overhead in service interface, or exceptions in computing resources, etc. One thing clear is that the performance of business process layer depends on the lower layer and measuring the performance of business layer includes indirectly measuring the performance of all the lower layers. But, most of the test frameworks developed focus on measuring SOA messaging performance, as we present in chapter 2. They almost adopt batch-style testing where all the test cases are executed in a sequence.

OMG recommended a standard SOA reference model, MDA (Model Driven Architecture) [1]. It is widely adopted in real world because it is normative and enables SOA system to be implemented in a business-centric approach. In the MDA, a business process is designed firstly in a way for satisfying business requirements and later services are bounded to the activities in the business process. Business processes are described in a standardized language (e.g. WSBPEL) and they are executed generally on a business process management (BPM) system.

For testing SOA systems implemented according to the MDA reference model in business-centric way, test harness should have business process simulation functionality so that it can behave as BPM and test overall performance at the same time. This means that the test harness can execute business process, perform tests, and gather metric values. The performance of the business process layer with which users keep in contact usually depends on the summation of the performance of the other lower layers. Therefore, for testing SOA system, test cases describing business process activities should be prepared.

In SOA system, test assertions may help develop tighter test cases which could be used as an input for SOA test harness. Any ambiguities, contradictions and statements which require excessive resources for testing can be noted as they become apparent during test assertion creation. Test assertions should be reviewed and approved to improve both the quality and time-to-deployment of the test cases. Therefore, best results are achieved when assertions are developed in parallel with the test cases.

Test assertions provide a starting point for writing conformance test cases or interoperability test cases. They simplify the distribution of the test development effort between different organizations while maintaining consistent test quality. Test assertions improve confidence in the resulting test and provide a basis for coverage analysis.

In section 2, we present some related works. Section 3 provides the concept of test assertion. In section 4, we describe a test assertion model. Section 5 presents cases of the test assertion and section 6 shows complex predicates of test assertions. Conclusions are presented in the last section.

2 Related Works

This section presents some test frameworks and script languages developed or proposed for SOA system.

Web Services Quality Management System

This system has been developed by NIA in order to measure Web services' quality on the criteria of WSQM (Web Services Quality Model) quality factors [2]: interoperability, security, manageability, performance, business processing capability, and business process quality. This system contributes to consolidate the quality factors of SOA. However, it requires expanding its architecture to apply SOA, system, because it targets only to Web services system.

ebXML Test Framework

This framework has been implemented by NIST and KorBIT for testing ebXML system according to OASIS IIC Specification [3]. It could test packaging, security, reliability, and transport protocol of ebXML messaging system implemented by ebMS specification [4]. The main purpose of this framework is to test conformance and interoperability of ebXML messaging system, and it is not proper for testing service oriented systems. Besides, it cannot test ad hoc status resulting from various events, because it is not event-driven but batch-style test framework.

JXUnit and JXU

JXUnit [5] and JXU [6] is a general scripting system (XML based) for defining test suites and test cases aimed at general e-business application testing. Test steps are written as Java classes. There is neither built-in support for business process test nor support for the event-driven features. However, as a general test scripting platform that relies on a common programming language, this system could be used as an implementation platform for general e-business tests.

ATML (Automatic Test Mark-up Language)

In its requirements, ATML provides XML Schemata and support information that allows the exchange of diagnostic information between conforming software components applications [7]. The overall goal is to support loosely coupled open architectures that permit the use of advanced diagnostic reasoning and analytical applications. The objective of ATML is focusing on the representation and transfer of test artifacts: diagnostics, test configuration, test description, instruments, etc.

Test Choreography Languages

These are standards for specifying the orchestration of business processes and/or transactional collaborations between partners. Although a markup like XPDL [8] is very complete from a process definition and control viewpoint, it is lacking the event-centric design and event correlation / querying capability required by testing and monitoring exchanges. Also, a design choice has been here to use a very restricted set of control primitives, easy to implement and validate, sufficient for test cases of modest size. Other languages or mark-ups define somehow choreographies of messages and their properties: ebBP, WS-BPEL, WS-Choreography[9]. The general focus of these dialects is either the operational aspect of driving business process or business transactions, and/or the contractual aspect, but not monitoring and validation. Although they may express detailed conformance requirements, they fall short of

covering the various aspects of an exhaustive conformance check, e.g. the generation of intentional errors or simulation of uncommon behaviors. In addition, the focus of these languages is mainly on one layer of the choreography – they for instance, ignore lower-level message exchanges entailed by quality of service concerns such as reliability, or binding patterns with the transport layer.

3 Concept of Test Assertion

A test assertion is a testable or measurable expression for evaluating the adherence of an implementation (or part of it) to a normative statement in a specification.

A set of test assertions may be associated with a conformance clause in order to define more precisely what conformance entails. Test assertions lie between the specification and any suite of tests to be conducted to determine conformance (See Fig. 1). Such a test suite is typically comprised of a set of test cases. These test cases may be derived from test assertions which address the normative statements of the specification.

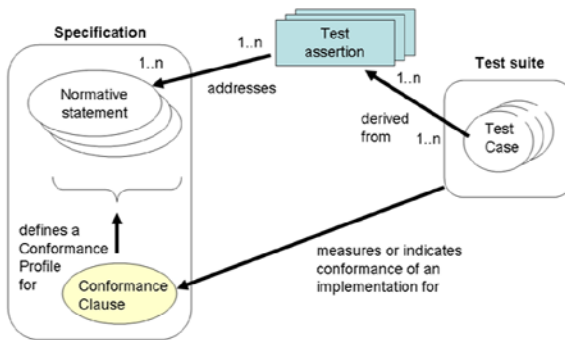


Fig. 1. Role of Test Assertion

Judging whether the test assertion is testable may require some knowledge about testing capabilities and resource constraints. Sometimes there is little knowledge of what actual testing conditions will be. In such cases the prime objective of writing test assertions is to provide a better understanding of what is expected from implementations, in order to fulfill the requirements. In other cases, the test assertions are designed to reflect a more precise knowledge of testing conditions. Such test assertions can more easily be used as a blueprint for test suites.

4 Test Assertion Model

This section aims to cover the simpler aspects of test assertions. Some more complex aspects are covered later in this section. Fig. 2 below shows the anatomy of a typical test assertion, and how its parts relate to the specification being addressed, as well as to the implementations under test. Some optional parts are not shown in the figure.

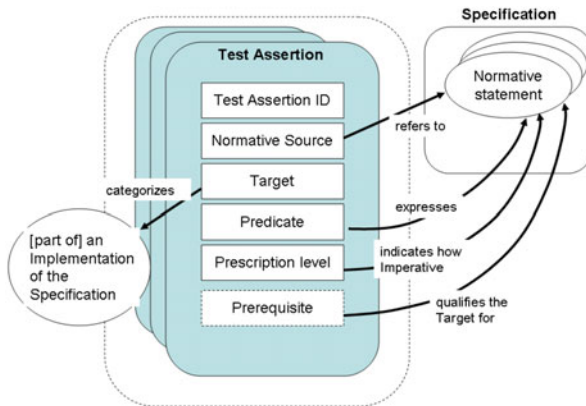


Fig. 2. General Anatomy of a Test Assertion

Some of the elements which comprise a test assertion are considered core while others are optional. A test assertion includes, implicitly or explicitly:

Identifier

A unique identifier of the test assertion facilitates tools development and the mapping of assertions to specification statements. It is recommended that the identifier be made universally unique.

Normative Sources

These refer to the precise specification requirements or normative statements that the test assertion addresses.

Target

The target categorizes an implementation or a part of an implementation of the referred specification, which is the main object of the test assertion and of its Normative Sources.

Predicate

A predicate asserts, in the form of an expression, the feature (a behavior or a property) described in the specification statement(s) referred by the Normative Sources. If the predicate is an expression which evaluates to “true” over the test assertion target, this means that the target exhibits this feature. “False” means the target does not exhibit this feature.

In addition, a test assertion may optionally include following components.

Description

This is an informal definition of the role of the test assertion with some optional details on some of its parts. This description must not alter the general meaning of the test assertion and its parts. This description may be used to annotate the test assertion with any information useful to its understanding. It does not need to be an exhaustive description of the test assertion.

Prescription Level

This is a keyword that indicates how imperative it is that the Normative Statement referred to in the Normative Source, be met. See possible keyword values in the Glossary.

Prerequisite

A test assertion Prerequisite is a logical expression (similar to a Predicate) which further qualifies the Target for undergoing the core test (expressed by the Predicate) that addresses the Normative Statement. It may include references to the outcome of other test assertions. If the Prerequisite evaluates to "false" then the Target instance is not qualified for evaluation by the Predicate.

Tags

Test assertions may be assigned 'tags' or 'keywords', which may in turn be given values. These tags provide you with an opportunity to categorize the test assertions. They enable you to group the test assertions; based on the type of test they assume or based on their target properties.

Variables

Test assertions may also include variables for convenience in storing values for reuse and shared use, as well as for parameterization.

As a test assertion has parts that can be evaluated over a Target instance (i.e. the Prerequisite and the Predicate), the following semantics apply to a test assertion:

- "Target not qualified": if the Prerequisite (if any) evaluates to "false" over a Target instance.
- "Normative statement fulfilled [by the Target]": if the Prerequisite (if any) evaluates to "true" over a Target instance, and the Predicate evaluates to "true".
- "Normative statement not fulfilled [by the Target]": if the Prerequisite (if any) evaluates to "true" over a Target instance, and the Predicate evaluates to "false".

5 Case Study of Test Assertion

Consider the following statement in the widget specification:

[requirement 101] “A widget of medium size **MUST** use exactly one AA battery encased in a battery holder.”

There are actually two requirements here that can be tested separately:

(requirement 101, part 1) A medium-size widget **MUST** use exactly one AA battery.

(requirement 101, part 2) A medium-size widget **MUST** have a battery holder encasing the battery.

Because of this it is possible to write two test assertions:

- **TA id:** widget-TA101-1a
Normative Source: specification requirement 101, part 1 Target: medium-size widget

- Predicate:** [the widget] uses exactly one AA battery. Prescription Level: mandatory
- and
- **TA id:** widget-TA101-1b
 - Normative Source:** specification requirement 101, part 2 Target: medium-size widget
 - Predicate:** [the widget] has a battery holder encasing the battery. Prescription Level: mandatory

The granularity of a test assertion is a matter of judgment. A single test assertion instead of two can be written here, with the predicate: “[the widget] uses exactly one AA battery AND has a battery holder, encasing the battery”.

This choice may later have an impact on the outcome of a test suite written to verify the conformance of widgets. With a single test assertion, a test case derived from this test assertion will not be expected to distinguish between the two failure cases. Using two test assertions - one for each sub-requirement - will ensure that a test suite can assess and report independently about the fulfillment of each sub-requirement. Other considerations such as the different nature of tests implied or the reuse of a test assertion in different conformance profiles [VAR], may also lead to the adoption of “fine-grained” instead of “coarse-grained” test assertions. Usage considerations will dictate the best choice.

6 Complex Predicates

Recall the previous example of [requirement 101]. The target can be defined as “a medium-size widget” or as just “a widget”. The latter is a natural decision if the specification requirement uses the wording: “[requirement 101] If a widget is medium size, then it MUST use exactly one AA battery and be encased in a battery holder.” For the simplicity of this example, if the two test assertion predicates for widget-TA101-1a and widget-TA101-1b are combined into one example, one possible outcome is:

- TA id:** widget-TA101-2a
- Normative Source:** requirement 101 Target: widget
- Predicate:** if [the widget] is medium-size, then [the widget] uses exactly one AA battery AND the battery is encased in a battery holder.
- Prescription Level:** mandatory

The target category is broad, but the predicate part is really of interest only for a subset of this category (the medium-size widgets). Usage considerations should again drive the decision here: a test suite that is designed to verify all widgets, and does not assume a prior categorization of these into small / medium / large sizes, would be improved with test assertions that only use “widget” as the target, such as widget-TA101-2a.

A test assertion predicate may, then, be a Boolean expression - a composition of atomic predicates using logical operators AND, OR, NOT. A test assertion predicate may also be of the kind: “if (condition) then (expression)”.

The predicate is worded in an abstract way, still close to the wording of the specification. No indication of what kind of test procedure will be used, such as how to determine the number and type of batteries, is given. Detailed criteria for the condition evaluation, such as what kind of battery holder is acceptable, are also not provided. These details are normally left to the test cases that can be derived from the test assertions. These test cases will determine the precise criteria for conforming to the specification. However, if a precise criterion for interpreting the battery holder requirement is provided in an external specification, either referred to directly by the widget specification or by a related conformance clause, then a test assertion must use this criterion in its predicate. Such a test assertion must then refer not only to the specification requirement in its reference property, but also to the external specification or to the conformance clause that refers to this specification.

Another case where a predicate is more complex is when its conditional expression involves more than one part of an implementation (or implementations). In some cases it is clear which one of these objects must be considered the target, while others are just accessory objects. Consider the following predicate: "the [widget price tag] matches the price assigned to the widget in its [catalog entry]", where price tags and catalog entries are both items that must follow the store policy (in effect the specification). In this case it may be reasonably assumed that the "catalog" content is authoritative over the price tag. The price tag can then be considered as the test target, while the accessory object may be identified by a variable which is then used in the predicate.

7 Conclusion

We presented a SOA test assertion model, which facilitates to make test cases in normalized form. In the model, we devised the concept of prescription level and normalized prerequisite for preparing test cases. By the concepts, test assertion can be transformed into test cases without semantic distortion. The model minimizes the human intervention in preparing test cases by automating some processes for translating test assertions into test cases. We showed two cases of complex predicates. Further studies are required to develop a test framework to check out automatically that test cases are conformed to test assertions.

References

1. Miller, J., Mukerji, J.: MDA Guide Version 1.0.1. OMG (June 2003), <http://www.omg.org/docs/omg/03-06-01.pdf>
2. Lee, Y., et al.: Web Services Quality Model 1.1. OASIS WSQM TC (October 2008)
3. Durand, J., et al.: ebXML Test Framework v1.0. OASIS IIC TC (October 2004)
4. Wenzel, P., et al.: ebXML Messaging Services 3.0. OASIS ebMS TC (July 2007)
5. Java XML Unit (JUnit), <http://jxunit.sourceforge.net>
6. JUnit, Java for Unit Test, <http://junit.sourceforge.net>
7. ATML, "Standard for Automatic Test Markup Language (ATML) for Exchanging Automatic Test Equipment and Test Information via XML". IEEE (December 2006)
8. XPDL: XML Process Definition Language (Workflow Management Coalition) Document Number WPMC-TC-1025: Version 1.14 (October 3, 2005)
9. OASIS, Business Process Specification Schema 1.0.1, May 2001 and ebBP, v2.0.4 (October 2006)

Idea Sketch of Application Service of e-Business Card and NFC Based on Mobile

Sun Kuh Noh¹, Byung Rae Cha², Dong You Choi³, Jung-Jin Park³,
Guen Sik Kim³, and Yoo Kang Ji^{4,*}

¹ Dept. of Radio Mobile Communications Engineering, Honam University,
#330 Eodeungno, Gwangsan-gu, Gwangju, 506-714, Korea
nsk7070@honam.ac.kr

² SCENT Center, GIST, Gwangju, Korea
brcha@nm.gist.ac.kr

³ Dept. of Information & Communications Engineering, Chosun University
#375 Seosuk-dong, Dong-gu, Gwangju, 501-759, Korea

⁴ Dept. of Information & Communications Engineering, DongShin University
#252 Deaho-dong, Naju, Jeonnam, 520-220, Korea
neobacje@gmail.com

Abstract. Today, all of the computing resources, environments of computing, and online application services have been moved to a computing environment based on mobile. In the future, especially mobile phone, can be a strong device not only supporting computing but being closely related to personal privacy. In this article, we sketch some ideas for various application services using e-Business card and NFC of objects on mobile environment.

Keywords: Component, e-Business Card, Application Service, Mobile Device.

1 Introduction

The existing computing environments are developing into a new paradigm of computing that convert to devices based on mobile - ubiquitous computing, cloud computing, and mobile environment etc. Especially, Mobile environment is a concept that means not 'fixed' but 'mobile(movable)' environment, more specifically, it means the hardware, software, network and service environment related to in-car devices, laptop, mobile phone, personal digital assistant (PDA). Today, all of computing environments, resources, on-line application services are moving towards computing environment based on mobile. In the future, especially the mobile phone, one of the communication devices, can be a strong device not only supporting computing but being closely related to personal privacy. We try to design the various application services by using e-Business card on mobile environment.

* Corresponding author.

2 Related Work

The smart-phone that both Samsung Electronics and LG Electronics recently rolled out, being embedded with business card scanner, can scan the business card, after then management changed it into e-Business card through image and character recognition technology. And now, many e-Business card softwares are being developed based on Android and iPhone. More developed, various types of e-Business card are followed to electronic business-card [1-2] of MS's outlook, Poken [3-4] of social network, and vCard [5-6]. First, Poken is a social business card being developed by Corp. Poken S. A. in Switzerland. It can transmit other's profile information once contacting with each other's doll-shaped devices containing USB, and also manage information using computer. The Korea Wireless Network (www.korwin.co.kr), a firm specialized in mobile service providers, developed a protocol stack (brand name: KorwinSTACK), which was supporting Bluetooth spec version 1.1, and implemented the application system for e-Business card (brand name: vCard++) by using it. Through this e-Business card system, business man will be able to exchange their photos, corporate logos, personal information by using mobile phone and PDA with Bluetooth in the future, instead of exchanging paper business card. Also, it is expected to be comfortable with exchanging own information not only to the business man but to the others.



Fig. 1. Poken of social network business card

3 Design of e-Business Card Application Service

We will design an e-Business card application service and NFC (Near-field communications) service of objects in mobile environments. The e-Business card supports information communications among users, stores and things using private communication network of Bluetooth, Wi-Fi, and so on as shown in Fig 2.

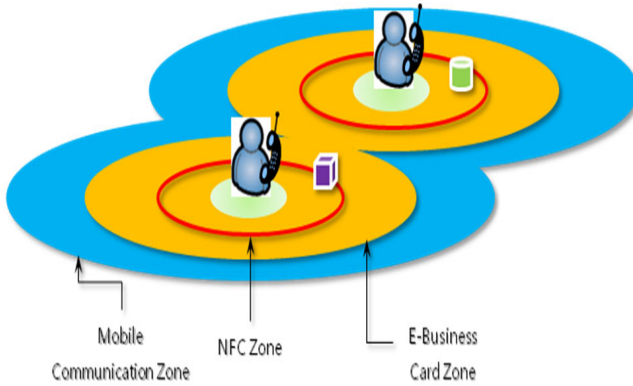


Fig. 2. Service zones of e-Business card and NFC using mobile devices

3.1 Data Structure of e-Business Card

The e-Business card contains <BusinessCard>, <STYLE> and the image files. The tag of <BusinessCard> and <STYLE> are formed with XML, each express personal information and graphic design on e-Business card as shown in Fig 3.

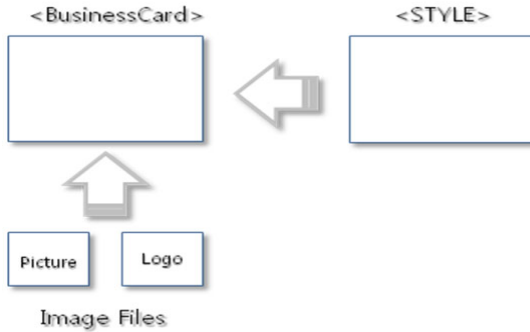


Fig. 3. Data structure of the proposed e-Business card

3.2 Application Service of e-Business Card

The application fields of e-Business card based on mobiles are divided into 5 parts:

- 1) Exchanging e-Business card among users,
- 2) Exchanging e-Business card between stores and users,
- 3) Exchanging e-Business card between automobile navigation system,
- 4) Exchanging information between Social Networks,
- 5) Key service of U-home.

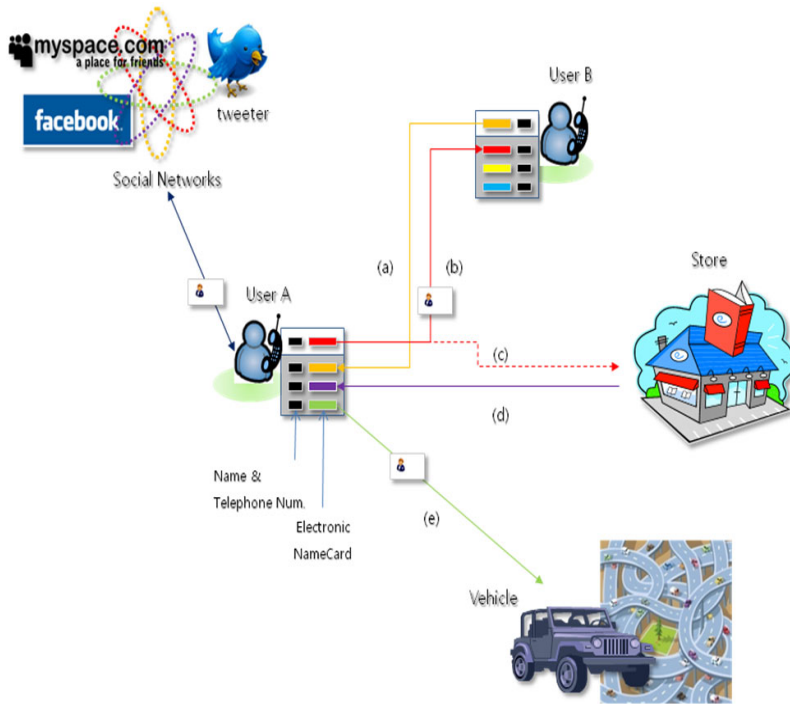


Fig. 4. Application services of e-Business card

Management of e-Business Card between Users and Social Networks. Many modern people are having the mobile device, and especially it becomes a necessity for the business man. The mobile device is performing electronic passport like text messaging, e-Mail, games, memo, scheduler, and etc, developed from just its means of communications. It is possible to exchange e-Business card among the mobile devices via Wi-Fi or Bluetooth communications rather than traditional exchanging the paper business cards among users. The exchanged e-Business card by users and social networks can provide an intelligent searching function, and is managed efficiently by Database Management System being embedded with the mobile device as shown in Fig 5 and Fig 6.

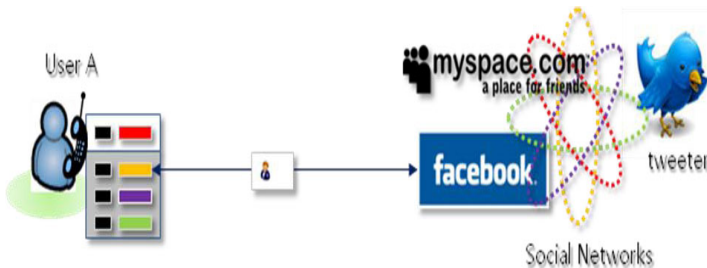


Fig. 5. Exchanging of e-Business card between user and social networks

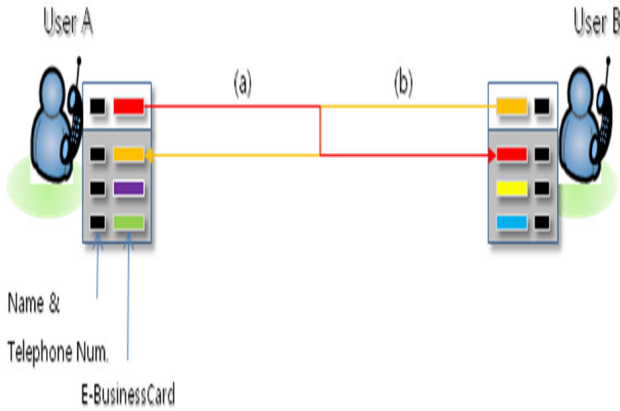


Fig. 6. Exchanging of e-Business card between users

Sales Management. The e-Business card supports the work collecting data and classification of customer management to business man. Especially, it is usable to search items for supporting enterprise-wide works around customers' profile as shown in Fig 7.

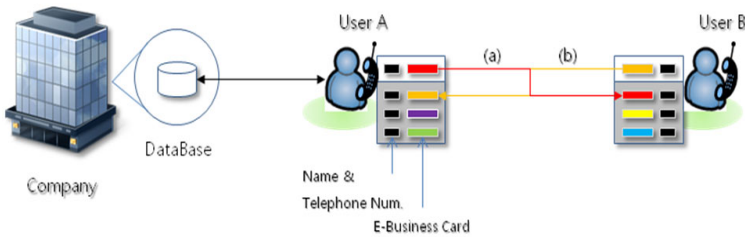


Fig. 7. Management of human profiles using e-Business card

Role of e-Business Card at Store. It is possible to exchange the e-Business card between mobile devices of its users, and exchange the information about their business at the restaurant or store, as well. The user can provide e-Business card about not all of information for exchanging and searching but necessary information because of privacy protection aspect as shown in Fig 8 (c).

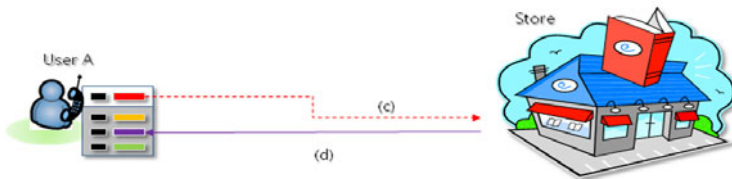


Fig. 8. Transmission and reception of e-Business card between users and stores

Services of Item Information and e-coupon on Mobile Internet Environment. It is able to provide the e-coupon service and search information about menu at restaurant or products sold at store using the information of store and restaurant listed on the e-Business card via mobile internet environment.

Privacy Policy. It is possible to set up the function about transmission and reception of e-Business card based on privacy policy and value of information. The service providers want to promote their information and get the customer's information, but the customers may regard this as unwelcomed things because of the privacy problem. So, it is needed to limit transmission of personal information and set up of the reception function.

The Security Function between e-Business Card and Setting up Destination on the Navigation in Automobile Smart System. The automobile smart systems are mainly divided into two parts: automobile smart-key and automobile navigation system. It is able to provide the function of automobile smart key adding up the message digest function to the e-Business card. It is possible that the function of automobile smart key supports identification technology of knowledge-base and ownership-base. The opening & shutting function of automobile smart key is based on hash value combining mobile phone number, information of e-Business card, private key, and its used record is remaining. The exchange between automobile navigation system and e-Business card is done for making users convenient by consisting ad hoc network between automobile navigation and mobile device, transmitting the information of destination rather than inputting search contents, searching and choosing. The automobile navigation sets up the destination, searching received information as shown in Fig 9.

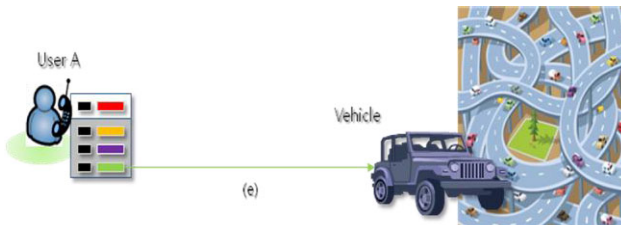


Fig. 9. Navigation setup of e-Business card

3.3 NFC Service of Objects

Near-field communications (NFC) promises to create a whole new paradigm for the vast majority of cell phone users and is emerging as a near-term reality. What has been described as the confluence of RFID and cellular telephony will bring with it a wealth of new applications [7]. NFC operates on the same RF principle as proximity cards. There are several advantages of integrating the solution into the cell phone. One obvious advantage is the “Swiss Army knife” approach to centralizing a user’s daily life (which also has its drawbacks that must be addressed in the design). Another

is that the phone affords a more capable engine than a card for enabling higher-level functions that demand greater memory and processing. Third, and most powerful, is that it provides a backend connection to the cellular network for high-level operations such as real-time loading of funds, security management, and telephone and Internet connections driven by inputs received from NFC interactions [7]. The evolution of NFC is based on an RF, data link, and MAC standard ISO/IEC 18092 [8]. The earlier proximity cards that are used to gain entry to buildings and mass transit in many countries are based on a decade old predecessor, ISO/IEC 14443 [9]. The 106 kb/s passive mode of the NFC standard is compatible with only part A of the predecessor, but since the application space is similar and they use the same RF circuits at the same carrier frequency, they are close enough that multiprotocol NFC front-ends can include operation for all. This is an important so that the phones can emulate transportation and identity cards already in use in very large numbers in part of the world. The multiprotocol operation simply does round-robin sampling for the presence of a response from one of the protocols, a process that occurs in less than a couple of hundred milliseconds. NFC and proximity cards are limited in effective range to just a few centimeters. This distance is sufficient so that you do not have to make contact, but some products are designed to require a slight touch to activate the link. The method of communicating is reactive near-field magnetic coupling at 13.56 MHz, which is a frequency available in all regulatory regions. The wavelength at this frequency is over 22 m, so the use of small loop antennas allows good magnetic coupling at short range while having very poor radiation efficiency. This combination is useful to both keep the communications private and be sure that when you point your wand, there is no ambiguity in your intent [7].

Secure key Function of U-home. Because Bluetooth and Wi-Fi have longer range connections, it is necessary to enter information to isolate and protect your own communications from others in the application area. And there is no way to indicate intent without typing it in. By touching two NFC enabled devices (e.g., a mobile device and an electronic door knob) as shown in Fig 10, we can immediately establish a Bluetooth and Wi-Fi connection between them. The intent of wanting the connection is established by the physical touch, and then the high data rate connection takes over. NFC is used as an enabler for a connection that was not previously established. Bluetooth and Wi-Fi already anticipate NFC as a pairing mechanism.

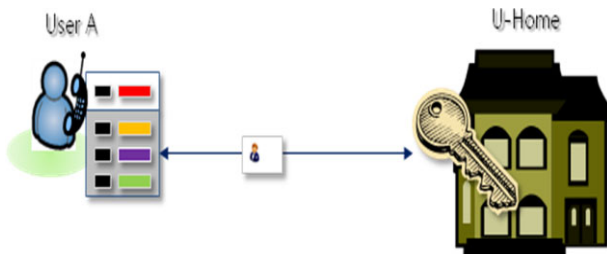


Fig. 10. Secure key function of U-home by NFC

The secure key service by NFC is provided adding up the function of message digest to electronic key. Especially, the function of secure key U-Home can support identification and authentication technology of knowledge-base and ownership-base. The opening & shutting function of electronic key is based on hash value combining mobile phone number, information of e-Business card, private key, and used record is remaining in mobile device by NFC.

The length of hash key for secure key service is proportional to endurable time against burst attack in secure aspect. The more long the hashed-key, the stronger on secure and it can provide much stronger key to electronic key U-Home, departing from existing electronic key that consist of combination of some numbers.

4 Access Control of NFC and e-Business Card

The access control of NFC and e-Business card mainly consists of four parts: DOM Plane, Security Plane, Interface Plane, Device Plane, as follows Fig 11.

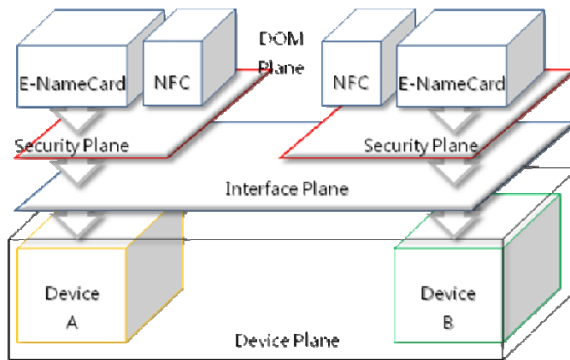


Fig. 11. Access control model of e-Business card

The access control of DOM Plane runs having logical relation between two cases: it runs in DOM aspect in one case of making and modifying of e-Business card and NFC objects or needing to convert after version upgrade, and in SAX aspect in the other case of composing menu and searching. The access control of Security Plane is logically related with access control and management about the e-Business card on the repository. The e-Business card is managed by the security level of itself (whether including digital certificate or not) and security level based on security policy of the user. The access control of Interface Plane is logically related with the interface between the version upgrade of e-Business card and the device. It supports interface for communication between device and the external to the multichannel. The access control of Device Plane is logically related with physical access about the storage device like USIM or external USB; it is for practical, physical access control.

5 Conclusion

All of the computing resources and environments of computing, and online application services are moving to a computing environment based on mobile. In the future, especially mobile phone, will be a strong device not only supporting computing but being closely related to personal privacy. In this article, we sketch some ideas for various application services by using e-Business card and NFC of objects on mobile environment. Especially, we introduce about the simple data structure and the outline of application service using e-Business card. And, we propose access control model for supporting the security function of e-Business card and NFC.

References

1. Bowman, M., Debray, S.K., Peterson, L.L.: Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.* 15(5), 795–825 (1993)
2. MS's Business card of Outlook, <http://office.microsoft.com/ko-kr/outlook-help/HA010068636.aspx>
3. Ding, W., Marchionini, G.: A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park (1997)
4. <http://www.poken.com/>
5. Fröhlich, B., Plate, J.: The cubic mouse: a new device for three-dimensional input. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2000*, April 01-06, pp. 526–531. ACM, New York (2000)
6. <http://en.wikipedia.org/wiki/VCard>
7. Fischer, J.: NFC in cell phones-The new paradigm for an interactive world
8. http://webstore.iec.ch/preview/info_isoiec18092%7Bed1.0%7Den.pdf
9. http://en.wikipedia.org/wiki/ISO/IEC_14443

Framework of Design Patent Map Systems Based on Dissimilarity from Populace Survey*

Chao-Chun Chen¹, Rain Chen^{2,**}, and Ting-Ting Dai³

¹ Institute of Manufacturing Information and Systems
National Cheng-Kung University, Tainan, Taiwan
chaochun@mail.ncku.edu.tw

² Department of Creative Product Design
Southern Taiwan University, Tainan, Taiwan
ocean@mail.stut.edu.tw

³ Department of Information Management
Southern Taiwan University, Tainan, Taiwan
daitingting7111@gmail.com

Abstract. Patent deployment has become a competition strength for companies. The intelligence property can keep the competition advantage of a company from opponents through the patent deployment which can be visualized by the patent map. Design patent map is an important strategic tool for establishing design strategies. Our past efforts studied the display techniques in design patent map, and the comparisons of design patents in United States and Taiwan. Of types of patents, design patents occupy a unique patent field, since design patents are not as definitive as other patent fields. Therefore, the construction of design patent map is extremely difficult. In this paper, we focus on the implementation of a design patent map system, which can be used to collect populace opinion on comparisons of the design patents. Our proposed framework for the design patent map systems consists of a dissimilarity collector, a patent database, a patent map visualization engine, a rendering tool, and a set of interface components. Hence, their functional integration fulfills both the populace survey and the automation of generating up-to-date patent maps. After prototype and demonstration, the proposed framework is able to offer a easy-to-use design patent map system to collect populace opinions and display design patent maps on web environments.

Keywords: design patent map, industry, populace survey, web technology, visualization.

* This work was partially supported by National Science Council of Taiwan and National Cheng-Kung University under grants NSC100-2221-E-006-267-MY2, NSC100-2221-E-218-049, and D100-36014.

** Corresponding author.

1 Introduction

With the arrival of knowledge economy era, enterprises are forced to treat intellectual property rights as efficient strategies in enhancing their competitiveness [1,2]. The World Intellectual Property Organization (WIPO) showed that over 90% of all inventions can be found in patent documents [3,4]. That is, the patent literature has been an enormous technical database for product designers or enterprises [5]. Hence, developing professional patent systems [6,7] and patent data analytical techniques [8] to interpret patent literature is one of most imperious demands for industries.

The United States Patent and Trademark Office (USPTO) offer three types of patents, including utility patents, design patents, and plant patents [9]. In this work, we focus on design patents [10,11,12]. Design patents tend to be the protection of forms and appearances, thus, comparison of design patents is more difficult than that of invention patents. The reason is that invention patents tend to have objective data for determination, such as method or process; however, design patents tend to be subjective determination, such as creativity and homogeneity which is still difficult to be evaluated or differentiated by current computer techniques.

Following the above mentioned property about design patents, we observe many judges on patent infringement are different from populace opinion in practice, because a judgement are decreed by only few justices. More and more patent experts believe in that collecting the populace opinion of patents as evidences will assist justices to make more fair judgements. However, no such system or platform assist barristers or patent experts to collect the statistics of populace opinion of patents so far. To achieve this goal, two challenges need to be conquered. The first is the convenience of access for populace. The system should be open any time and for everyone. Most users has low desire to make the patent comparisons on the dedicate system and even with specified time schedule. The second is the efficiency and scalability of visualization transformation.

In this paper, we consider the common features of tackling the above mentioned challenges in automated design patent map systems, including the collection of populace survey, automated visualization transformation, graphical display of dissimilarity of patents, and patent maintenance. Ideally, once users offer their comparison results among patents, the system would automatically generate the up-to-date design patent map while administrators or barristers request. The design consists of four main components, including the collection of populace survey, automated visualization transformation, graphical display of dissimilarity of patents, and patent maintenance. To achieve a proof of concept, we develop a prototype of a design patent map system based on our proposed system architecture. The prototype is built based on well-known Apache web platform with PHP and the database MySQL. The demonstration results presents the web-based interfaces that get higher user experience. Currently, the prototype is adopted in patent-related courses of Dept. of Creative Product Design, Southern Taiwan University, and will be promoted to industries in the future.

The rest paper is organized as follows. Section 2 describes the motivation and scenario design. Then, we present the system design in Section 3. Section 4 shows the demonstrations of our prototype. Finally, we give remarks and conclude the paper in Section 5.

2 Architecture and Scenario Design

Figure 1 instantiates the user scenarios according to our condensed practical experience in design patent map systems. The first component, *patent maintenance*, is used to offer CRUD (creat/retrieve/update/delete) operations on patents' information for administrators. Besides, administrators can also retrieve the current status of patent comparison, e.g., the number of participants and the dissimilarity value of a given patent pair.

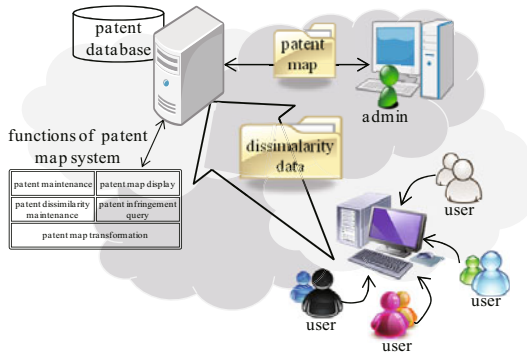


Fig. 1. Reference architecture for populace-used patent map systems

The second component, *patent map display*, is used to display a illustrative figure of a design patent map which is generated by patent map transformation discussed later. The patent map display component mainly includes a set of functions which drawing the patents' locations of a design patent map into a web page. Whereon, we further develop additional functions that can display nation/enterprise and basic information of patents by just mouse movement for users.

The third component, *patent dissimilarity comparison*, would guide users to compare the dissimilarity of patents. Assume that a number of users are invited or informed to offer their opinions on patent comparisons through the social relationship. Our system activates permanently, thus, every user can compare the dissimilarity of patents whenever he or she is available to make the survey.

The fourth component, *patent map transformation*, would automatically transform the dissimilarity of patents into a patent map. The kernal method of

the component is based on multi-dimensional scaling which is a traditional problem in computational statistics. We implement two versions in the component: one is the Kruskal method, and the other is the genetic algorithm-based method. Due to the content of the two methods is beyond the scope of this paper, interesting readers can refer the details to the related papers.

The fourth component, *patent infringement query*, allows users to query whether the given patent pair could incur infringement of intellectual property law. The component can return a deterministic result if the judgement record of the patent pair exists in the patent database. Otherwise, the component counts the similar cases to the given patent pair in the design patent map, and then returns a probabilistic result to users. The infringement query processing is an application of the design patent map. Enterprises can make many competitive business decisions based on such the application. Finally, the last component, *patent database*, offers the storage of all data used in the above mentioned components.

3 System Design

We present the system design in two levels: one is from the administrator aspect, the other is from the populace aspect. The function design presented in this section is based on our experience. In practical, modules in each aspect are very flexible. Interesting readers can easily modify our design to fit to different application needs.

3.1 Modules for Administrators

Modules for administrators include patent maintenance, patent map transformation, and patent map display. Figure 2 illustrates the action flowcharts of the modules, which depict the interaction between administrators and associated modules. The patent maintenance (PM) module offer administrators the CRUD operations on patents in the patent database. The patent maintenance module has two main functions: (1) obtaining requests of administrators and (2) manipulation in the patent database. The first function is to obtain the request command and associated patent information through interaction to administrators. After that, the second function is called to transform the request command and associated information to SQL statements and then send to the patent database to accomplish the administrators' request.

The patent map transformation (PMT) module is used to transform the dissimilarity to a two-dimensional map. Basically, the module is the most critical component of the automated patent map systems. We implement the PTM module by using the multi-dimensional scaling (MDS) technique [12].

The patent map display (PMD) module is used to display the two-dimensional map generated by the PMT module on a web page. The basic functions of this module include a set of plotting functions to generate an interactive figure.

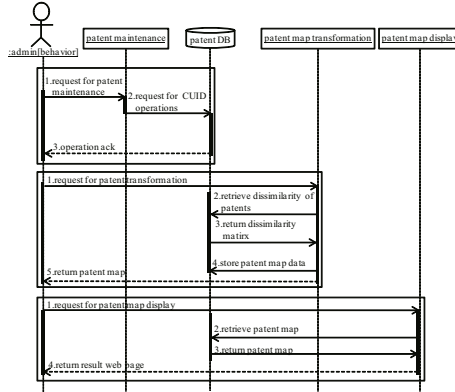


Fig. 2. Actions flowcharts of modules for administrators

3.2 Modules for Populace

Modules for populace include patent dissimilarity comparison and patent infringement query. Figure 3 illustrates the action flowcharts of the modules, which depict the interaction between populace and associated modules. The patent dissimilarity comparison (PDC) module is designed to guide users to compare the dissimilarity of all patent pair as possible. In practice, experts recommend two type of comparison view for users: one is the table view and the other is the matrix view. The table view shows the list of patents whose images are displayed for users. A user can then select any two patents for comparison. The matrix view shows the current comparison of patents in a matrix, so that a user can easily grasp the overall comparison status. Then, a user may select a pair for comparison if he/she finds that the number of participants who survey on the pair is insufficient.

The patent infringement query (PIQ) module can let users understand the probability that the user-specified patent pair violates the intellecture property

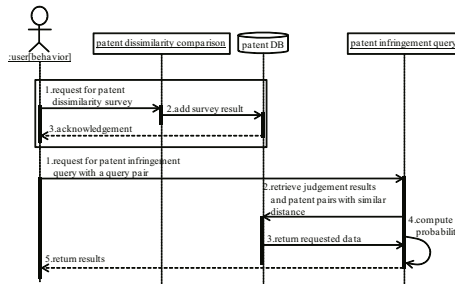


Fig. 3. Actions flowcharts of modules for populace

law. The PIQ module estimate the probability of the patent infringement based on the current comparison results from populace.

4 Demonstration and Performance Study

In this section, we first demonstrate our prototype, and then, show the primary version of our performance study.

4.1 Prototype Development

In order to offer convenient access and usage of a design patent map system, we develop the prototype by using web programming with a database system. The physical development platform include PHP 5.2.6 based on Apache server 2.2.8, and the patent database is built on MySQL 5.0.51a. These development tools are very popular and can be cross-platform, thus, our prototype can be deployed on almost any kinds of general PC platforms which have enough computation capacity, networking module, and support the above mentioned tools. Currently, our prototype can be accessed at <http://cis.mis.stut.edu.tw/PTMS/>.

4.2 Scenario Demonstration

The implementation reveals a real application that users can contribute their wisdom on comparing dissimilarity of patents through a ease use web interfaces. In addition, administrators can display the graphical representation of current dissimilarity of patents. Most manual settings, such as the parameters of the transformation engine, are hidden from users. Such the design is to simplify users' usage and focus all concerns on patent comparisons. Most management tasks about patents or system parameters are assigned to administrators. In the following, we present some screenshots to demonstrate the use scenarios.

Figure 4 depicts the patent dissimilarity survey from populace. Recall that there are two types of patent pair selections in our framework, including the table view and the matrix view. In the figure, we show the comparison process by through the two types of patent pair selections. The first subfigure is the screenshot of the patent pair selections by the table view. A user can select two patents for comparisons through the assistance of thumbnail patent photos, for example, patents 1 and 6 are selected in the first subfigures. Then, the prototype will guide the user to compare the dissimilarity of selected patent pair, as shown in the second subfigure. The third subfigure shows the screenshot that a user can determine to select a patent pair by either the table view or the matrix view. the forth subfigure shows the screenshot of the matrix view.

Figure 5 shows the patent map with a visual tool. The location data in the figure are from the patent transformation engine. From the patent map, patents that most participants deem they are similar are placed closely. For example, patents 11 and 5 are very close which means most participants deem they are very similar.



Fig. 4. Screenshots of the patent dissimilarity comparison

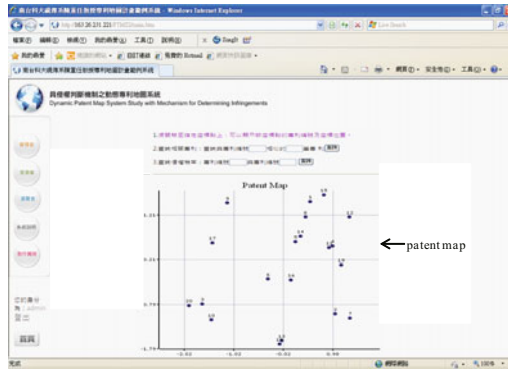


Fig. 5. Screenshots of the patent map display

5 Conclusions and Future Work

This paper has proposed the framework of a design patent map system for collecting the common opinion from populace survey. Based on the proposed framework, our efforts have developed modules for populace survey, patent map transformation, patent map display, and patent maintenance. All the modules are implemented on the web platform, so that users can use the system almost any time and anywhere, once they have internet access. Demonstrations have shown the various use scenarios on the prototype of the proposed framework. Incidentally, the prototype has adopted as a platform in the patent courses of Southern Taiwan University.

At present, we proceed to design the patent map transformation with higher efficiency. The patent map transformation module is the engine of a design patent map system. Hence, changing an efficient engine will significantly improve the performance of the system. In addition, we continue to revise the website design to increase use experience on our system [13]. Currently, our website design is based on the master thesis [14]. The human-server interface will be improved according to the students' feedback from Prof. Rain Chen's patent courses.

References

1. Chen, R.: Design patent supported strategy: The core for the enhancement of corporate competitiveness. National Digital Library of Theses and Dissertations in Taiwan, 122 (2004)
2. Chen, R.: Design patent map visualization display. *Expert Systems with Applications* 36, 12 (2009)
3. Liu, S.J., Shyu, J.: Strategic planning for technology development with patent analysis. *International Journal of Technology Management* 13, 19 (1997)
4. Chen, J.L., Liu, S.J., Tseng, C.H.: Technological innovation and strategy adaptation in the product life cycle. *Technology Management: Strategies and Application* 5, 19 (2000)
5. Herce, J.L.: Wipo patent information service for developing countries. *World Patent Information* 23, 13 (2001)
6. Huang, S.H., Ke, H.R., Yang, W.P.: Structure clustering for chinese patent documents. *Expert Systems with Applications* 34, 7 (2008)
7. Wanner, L., Ricardo, B.Y., Brugmann, S., Codina, J., Diallo, B., Escorsa, E., Giereth, M., Kompatsiaris, Y., Papadopoulos, S., Pianta, E., Piella, G., Puhlmann, I., Rao, G., Rotard, M., Schoester, P., Serafini, L., Zervaki, V.: Towards content-oriented patent document processing. *World Patent Information* 30, 13 (2008)
8. Tseng, Y.H., Lin, C.J., Lin, Y.I.: Text mining techniques for patent analysis. *Information Processing and Management* 43, 32 (2007)
9. USPTO United States Patent and Trademark Office: US patent law (2008), <http://www.uspto.gov>
10. Chen, A., Chen, R.: Design patent map: An innovative measure for corporative design strategies. *Engineering Management Journal* 19, 15 (2007)
11. Chen, A., Chen, R.: A seven-track protection mechanism for design knowledge. *IIGSS Academic Publisher Scientific Inquiry* 8, 11 (2007)
12. Kruskal, J.: Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29, 15 (1964)
13. Shneiderman, B.: The next generation of graphical user interfaces: Information visualization and better window management. *Displays* 17, 5 (1997)
14. Yao, B.B.: The research on patent map system and interface planning. National Digital Library of Theses and Dissertations in Taiwan, 95 (2010)

Service Quality Assessment in Provision of Library Services

Masitah Ahmad¹, Jemal Abawajy¹, and Tai-hoon Kim²

¹ School of information Technology, Deakin University, Victoria, Australia
{mahma, jemal}@deakin.edu.au

² Department of Multimedia Engineering Hannam University, Daejeon, Korea
taihoonn@empal.com

Abstract. Libraries are dynamic organization providing dynamic services to their patrons. However, advances in IT and networking have created a serious competition to libraries from many information service provisioning agents. The purpose of this study is to understand the issue relevant to service quality assessment in libraries and subsequently develop a conceptual model of aligning library quality of service assessment for effective reporting of library value and performance to stakeholders. In this paper, we will focus on the provision of library services and the nature of the relationship and interactions between various components of library services.

Keywords: Service Quality Assessment, Provision, Library Services.

1 Introduction

Libraries have changed the way they provide services. Today's libraries mostly offer their services in electronic and web-based form. These services include online databases, electronic serial and digital documents [1]. Therefore, quality of service (QoS) is one of the most important factors and is a major issue on the research agenda for electronic services [2],[3]. Commercial information service providers are now competing in the information marketplace. Thus, libraries are facing stiff competition from many information providing agencies. Also, libraries are being held accountable from agencies such as their parent institutions and accrediting bodies regarding the quality and the impact of the service they provide to their community of service users. Unfortunately, the traditional quantitative measurement of library collections is no longer appropriate, nor applicable, as a means of quality of service assessment.

Currently, different library use different model to measure their QoS [4]. Traditional measures of library services, such as counting interlibrary loans, circulation statistics or percentages of reference questions answered correctly have become obsolete [5]. The major issues with current methods is that performance indicators only show the raw materials or potential input, not the results our output. In many service industries, companies have created programs which incorporate a survey that elicits customers' assessments of service quality, plus a feedback loop through which service changes are implemented and then evaluated with subsequent

survey data. However, the determinants and measurements of service quality in library become unique compared to other service sectors. This is because QoS in libraries can have many aspects including (1) user-orientation of services, (2) accuracy and reliability of the services, (3) speed and currency of the services, (4) accessibility, (5) competence of staff; and (6) effectiveness and efficiency [6].

Thus, in the age of information-rich era and cost-effective information providing agents competing with the libraries, library managements need to change the way they measure QoS and build a climate of continuous improvement in all areas of library service. Especially work directly with customers to deliver such service on a continuous basis. Customer satisfaction is seen as the libraries highest priority. Due to that, the successful in libraries services will be achieved if customers are satisfied. On the other hand, the services that been used by the user been pay based on QoS. Library should centre on service provision and improvement on building a good relationship between stakeholders with their services.

In this article, we aim to understand the issue relevant to service quality assessment processes in libraries and subsequently develop a conceptual model of aligning library quality of service assessment for effective reporting of library value and performance to stakeholders with focus on service quality and factors that exert significant relationship in regards to the QoS in libraries.

The rest of the paper is structured as follow: first we present layered architecture of the library system. Next, we purpose a conceptual framework for assessing QoS in provisioning library services. Finally, we concluded the paper in the last section.

2 Overview of the Problem

The objectives of e-government are to offer efficient, high quality administrative on-line services to citizens and businesses. They also include an effort to streamline government's internal processes to improve Quality of Services (QoS), reduce costs and increase productivity [7]. In other words, the primary objectives are to bring dramatic improvements in the quality of government's interactions with its. Due to that, all organization which involve with the services include libraries must deliver a benefit services to their customers. Libraries have to continually engage their users to continue using and value library resources, services and facilities. With the dramatic increase of available materials and user expectations, libraries are forced to exploit new technologies to fulfil their missions with relatively limited resources [8].

Libraries are dynamic organization providing dynamic services to their patrons. However, advances in IT and networking have created a serious competition to libraries from many information service provisioning agents. Therefore, it is imperative for libraries to seek ways to ensure that their services meet and preferably exceed stakeholders' expectations. Libraries are adapting service-oriented provisioning. Hence to that, the evaluation of the performance in library is one of the major concerns and an integral part of the library management. Perhaps could attract the new user to use their services.

Libraries organizational structures have changed considerably in just a few years. Long vertical hierarchies are disappearing and in their place are many structures that

recognize the new technological realities. Furthermore sets of performance measures and indicators (example *ISO 11620 1998, PROLIB-PI and IFLA Handbook*), only include measures mainly for traditional library services [9]. Due to that, it's important for libraries to come out with new performance indicators that cover electronic services in libraries. A good relationship between library and users will encourage users to come and use library resources. These relations are valued as the interactive activities. Repeated review of service goals and objective in relation to user expectation is the fundamental need in service quality in libraries. In addition, technology and competitors helps libraries to make the young generation more belief in the gathering, evaluation and the use of information using libraries. Based on this statement, library as service organization need to have a motivated staff that could give excellent service and directly work with the Stakeholders.

Stakeholders are those people inside and outside the library who have a vested interest in the library. In the current information age, libraries seeking to meet the information needs of their clients are increasingly looking to modern electronic technologies (including computing devices, mobile phones; and the Internet). The goal of a library is not to make a profit but to satisfy readers' needs for information. Libraries compete with other departments for financial resources. Due to that, financial support of a library would depend on the user satisfaction with its services. Unfortunately, there has been no consistent way for assessing the quality of services the libraries provide to end-users.

The purpose of this paper is to propose a conceptual framework for library institutions to better assess the quality of their services. We will focus on service quality and factors affecting it. The final complete framework will enable the libraries to identify any gaps in their service management and provision so that these requirements can be built into their business and service quality improvement planning, offering clarity of direction for service managers and transparency of development to meet business and client need.

3 Related Work

[2], using SERVQUAL as a tool to measure QoS in libraries (responsiveness, empathy, assurance, tangible; and reliability). The study focuses on user perception of university library services and their level of satisfaction with library services. The results show responsiveness ($m = 3.39$), followed by empathy ($m = 3.36$), assurance ($m = 3.31$), tangible ($m = 3.13$) and reliability ($m = 2.47$). Due to that, [5] suggests that although the academic staff are using the library services, their perception of the quality of library services is "average"

Alia and Kanwal [3] on the other hand, considered only five attributes (tangibles, reliability, responsiveness, assurance and empathy) in SERVQUAL for their study. This study reveals that user expectation was high as compared to their perceptions. Users had overall positive view of libraries service quality. Service quality was found good regarding library staff polite behavior and their ability to perform quickly.

However, in term of modern equipment, visually appealing facilities and knowledgeable staff still lacking [3].

SERVQUAL is the most popular and standard tool to measure quality for library services [10]. Using explorations method, [10] investigate the overall service quality of library system. Modified version of the SERVQUAL questionnaire for data collection was uses. Differences gap between perceived and desired services were calculated and ranked to indicate services meet, exceed or fall short of the user expectations. From this study, it revealed that all the attributes (assurance, collection and access, empathy, library as a place, reliability, responsiveness and tangibles) are desired service in higher than perceived service. It was concluded that the institution pay attention to printed resources, easy access to online resources, modern equipment and facilities.

Reza and Hossein [11], study the prioritizing academic library service quality indicators using fuzzy approach. Once again in this study, SERVQUAL model been used. [11] claimed the measurement of the dimension and component of SERVQUAL show criticized; (1) These method disregard the vagueness of the individual judgments and their value changes when converted to number; and (2) The evaluators' subjectivity, judgment, selection and priority would have a big impact on the results of the methods.

Kaur [12] also suggest different measuring instrument will be needed to be used, because now libraries using electronic services to their clients. A survey was conducted among faculty, graduate and undergraduate student using modified version of SERVQUAL done by [13]. Four factors found through this exploratory analysis was; (1). Affect of service (organization), (2). Collections and access, (3) Library as a place; and (4) Affect of service (personal).

[14] study the service quality of university libraries at two university libraries. Using questionnaire to measure this service quality with total of 400 students were interviewed. This study has been used SERVQUAL modified version as a tool to assess service quality and user satisfaction. As a result, [14] find that five library service quality factors (quality of library service provided, quality of information and library environment, reliability, quality of online catalog system and confidences) positively affected students' satisfaction and crucial to an excellent library system.

As conclusion from the related work done by previous researchers [2],[3],[10],[11],[12],[13],[14] we can conclude that SERVQUAL only measure QoS based on the certain services in libraries.

4 Reference Library System Model

Libraries provide a variety of different services to a variety of different user groups. Generally, provision of service, not profit, in which intangible benefits are provided to individuals, is the main objective of the libraries. In the context of this article libraries terms is used to indicate the full range of library services (Digital services – electronic services) and information services (Information technology) without particular on their formal title.

There are two categories for library user services. These include library public user services (refer to circulation, bibliographic instructions, distance learning, government documentation, reference, special collections and so on) and library technical services (focus on procedures and operations of maintaining, developing and supporting library collections behind the scene such as acquisition, cataloguing and classification, interlibrary loan and document delivery, serials, and systems) [5]. Normally internal services involve with service provider (SP) that supply the material and product for library. For example like book, magazine, database and so on. So each time when this product has been deliver to Public Services, the staff inside this department need to do a process that include i)Acquisition, ii)Cataloguing, iii)Archiving and iv)Retrieving. After the process been done, the user can be use the services that known as external services. Loan and reading facilities is the services that can be use by the user.

The library services can be generally classified as technology-driven services (i.e., an Internet accessible collection of services) and non-technology based services. An example of the later is the information product (i.e., the content and quality of the information) and the service components which include the facilities and the computerized and human assistance that deliver the information product to its users) [15]. The trends in library service provisioning indicate that more and more library processes and workflows will take the form of automated systems built by combining a variety of services.

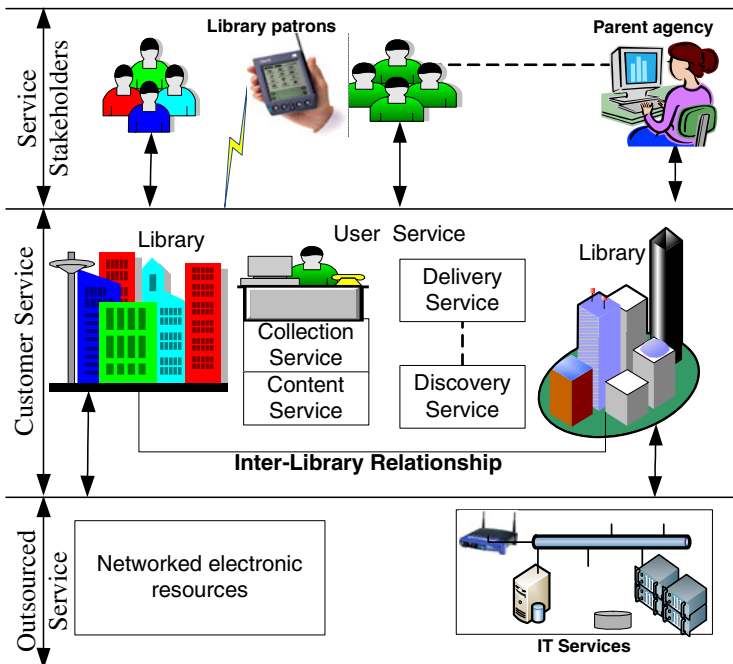


Fig. 1. Library service, their users, and the interactions between them

Library patrons are parts of other systems (researchers, disciplines, families, occupational groups). The library service usage behaviour of the library patrons can be expected to range over numerous other systems in addition to the library service (e.g., colleagues, newspapers, bookshops, other libraries, meetings, etc.). A given library service will be only one part of any given user's range of information. Also, the expectations of the library patrons and the parenting agencies are quite different. The parent agency, for example, requires the libraries to provide services in line with their mission and goals [16].

Basically, most of this result may have different perception for each of the stakeholder group. Because, the user may look library quality in different angle based on their experience with the services that they use. Some of them may not care for the efficiency on the locale process, but more particular on delivery of service effectiveness. It is quite different with the funding or parent institution where there are more interested in the benefit that library can give to them and in the library's cost effectiveness. The libraries need to prove their value and contribution to the parent agency whole, worthy of their share of the agency budget. From the parent agencies prospective, libraries are expected to provide more efficient ways of delivering services to achieve greater returns on investments. Thus, librarians should fully understand and aware of the supporting agency goals and strategic directions, since library operations are sometimes seen as peripheral to an agency's core competencies. They must align themselves with agency goals, and make sure the value of the library's contributions is known throughout the organization.

Libraries not only have their own internal systems and procedures (e.g., collection service, delivery service and discovery services), but also participate wider sorts of systems (e.g., networked electronic resources refer to a database of Internet network service) and with other libraries (i.e., libraries depend significantly on other libraries for interlibrary loans and, more recently, on the collective use of computerized bibliographical utilities).

From Figure 1, it can be deduced that library service is a bundle of services including. Thus, within the realm of library, services are comprised of multiple components and each component may have its own unique result of an outcome assessment. Most library services are bundles of core, facilitating and supporting service. Moreover, these services are multi-dimensional. For example, airline service includes a core service, transportation, plus check-in service, in-flight meal service and so forth. Thus, assessment of a service can be decomposed into a series of interrelated stages: assessments of performance, service quality and value.

Digital revolution also has made library collection more visible, accessible and usable. The amount of digital information used in library largely network-based. As a result of advance in IT add networking. The prevalent format, the speed of information creation, delivery and dissemination and user need and expectations have changed. They also have been using IT to computerize their technical processes develop networks and provide innovative, build databases, range of administrative and using intelligent information services [6].

5 Multi-Stage Model of Service Quality

Although libraries have had tools (e.g. accreditation; Benchmarking; Total Quality management(TQM); International Standard (ISO11620); Services quality (SERVQUAL); and Library service quality (LibQUAL)) to assist them in decision making on the basis of tangible indicators such as size and quality of their collections, there have been few instrument for monitoring and measuring service quality. Research in library system has so far focused on the quality of the library in terms of its collection, the size of its holding as well as various counts of its use [5]. There are also researches on the performance of libraries in terms of its performance as viewed from the library patrons prospective. We argue that these approaches mainly measure library customer satisfaction and not service quality. One such tool is the SERVQUAL is the most popular and standard tool to measure quality among the customer [2],[3],[10],[11],[12],[13],[14]. However, elements in this SERVQUAL could not be use to measure the QoS which specific to e-services in the libraries. Unfortunately, they do not describe performance or indicate whether service quality is good, indifferent, or bad. The need to understand what library customers expect in terms of service quality is now necessary for good management.

The provision of library services depends on the provisions of other services such as IT resources and online databases. Therefore, the QoS in library context should be looked at as a multidimensional factor encompassing customer services as well as outsourced delivery of service effectiveness. The nature and the need of library patrons vary from one sort of context to another. Different sorts of inquiries call for different sorts of services. Therefore, an approach that takes into account these factors should be developed in order to assess quality of service provisioning in library environments with a sense of coherence—a sense of fitting together to form a whole. Also, libraries have been traditionally classified into academic, school, public, and special (i.e., specialized and usually in support of industry or public administration). The traditional types of library services are generally noticeably different from each other, although each type of library may vary from one country to another. Therefore, coming-up with a general framework for QoS assessment in library service provisioning is a complex undertaking.

The model presented in Figure 2 is an original represent of how the authors believe service quality assessment stage appear within a Libraries Services environment. The key characteristics of this model include: Stakeholders view (referred to Beneficial Outcome); Service delivery view (Utilization, Capability and Access Quality Indicator) and Resource view (IT Services, Digital Services and User Services). Figure 2 show various library services and their multiple constituencies with differing interests (e.g., patrons, sponsors, funding agency, etc.) as well as the interactions between them as a system.

Figure 2 also shows a multi-stage service quality assessment model that we propose in this paper. Information Technology (IT) is a service that libraries get from other department for example IT department which related with the IT facilities. Among those services includes photocopy machine, connection speed, level of difficulty in operating computer interface, collection type, record of collection, renewal and timeliness, success rate of information search and, level of difficulty of

search interface. Digital services, includes Online collections search (e.g. online catalogue search number; electronic periodical search; online periodical table of contents search; CD-ROM database search and transfer of electronic documents), and Online library services (e.g. Online reservation number; online overdue notices number; online applications for cooperative library services number; online information recommendations; E-mail assistance services; search library www information services and online reference consultation). Finally for user services, it's includes database (e.g. number of online abstract database; number of online full text data), OPAC (e.g. number of electronic periodicals; number of electronic books; number of electronic reference materials).

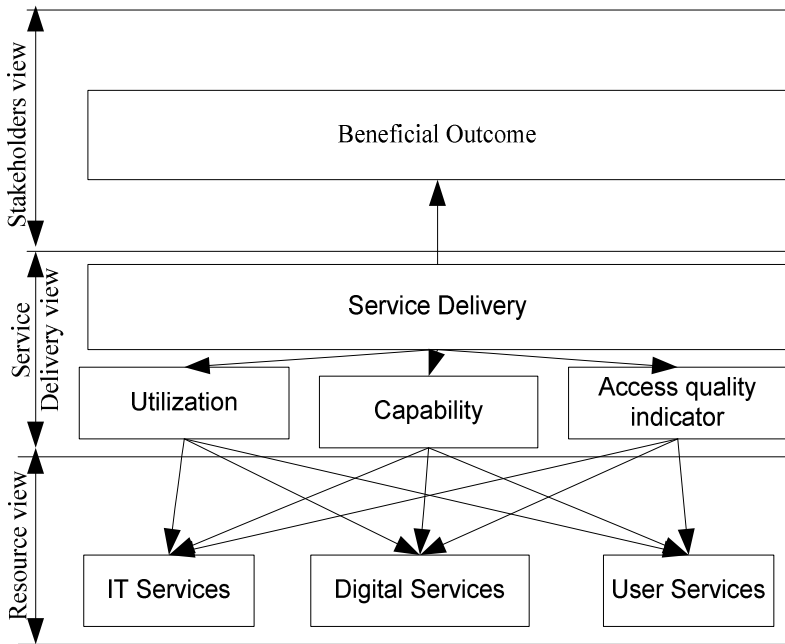


Fig. 2. Multi-stage service quality assesment model

Service delivery includes such as longer hours of service and better assistance as well as resources selected for their quality, authority, accessibility, currency and subject relevance. The type of capability being provided needs to be related to the sort of demand to be served. Immediate availability and related measures of document delivery are evidently measures of a library's capability. The utilization of the services is another major factor to be taken into the determination of library service quality provisions. The beneficial outcome is a multi-dimensional output. Its first dimension is technical quality, meaning the outcome of service performance. Its second dimension is functional quality, meaning subjective perceptions of how service is delivered. Functional quality reflects consumers' perceptions of their interactions with service providers. The third is program quality: the range of activity programs, operating time, and secondary services. The fourth is interaction quality, or outcome quality.

The intention of this study is to investigate and devolve a conceptual framework for quality of service assessment in the provision of library services. Currently, the services provided by libraries are inherently disjointed set of reference models. There must be a way to create useful relationships across the various services. The proposed conceptual framework aims to rectify this through the accumulation of reference models that provide a consistent view of library activities, and a shared means of expressing them. The framework would assist libraries in strategic planning. It would provide a tool to help them establish priorities, guide investment, and anticipate future needs in uncertain environments. Moreover, the framework will help develop standardization in quality of service assessment in libraries service provisioning. These standardized processes would facilitate consistent performance measures and benchmarks.

6 Conclusions and Future Directions

QoS is a crucial issue for each organization which involve with services. Due to fulfill the satisfaction and the need of their user, many organization have to advance in their services. Therefore using IT in the organization is one of the important ways has been formally recognized by the user in fulfill their need and satisfaction. Providing reliable QoS is one of the important requirements in Cloud computing environment. It is important to set out clearly the relationships between the Service Provider (SP) and the user.

Libraries also one of the important institutions which have services that can be given to their patron. Emerging phenomena like e-information services [17], e-library services which related with E-Facilities, E-collection; and E-Service has been done at most of the libraries. Unfortunatly, there is lack of authoritative guideline or criteria to help library define how QoS of a libraries performance should be provide. Especially, related with the proper measured and improved to serve end user better. Both sides need a benchmark-against which tests the efficiency and effectiveness of the services. The lack of understanding on both sides means the objectives and goals are not defined. As a result, the user is dissatisfied with the services. Especially work directly with customers to deliver such service on a continuous basis. They also had profound impact on the way library services operate. QoS has been become one of the most important factors and major issues on the research plan for electronic services in the digital age.

Number of indicators has been developed and tested in libraries. However, elements in this indicator could not be used to fulfill the measure the QoS which specific toe-services in the libraries. Libraries should continue to use these indicators to compare result and method generally consensus on validated the indicators. Libraries must be able to show whether they are developing in the right direction, their services has been accepted by user and whether their services has been offered in a cost-effective way [18]. Literature reveals that evaluating QoS in libraries should be considered to a greater extent. Due to that, more evaluation studies on the basis validated methods are needed.

In this paper, we have reviewed the progress that has been made in quality of service assessment in provisioning of library services. We find that comprehensive approach that encompasses various aspects of the library services as whole is lacking. This paper proposed a conceptual model applicable to all types of libraries to effectively assess quality of their service and increase the alignment of assessment and the library's ability to report value and impact. One research direction is to further extend and then test and confirm the proposed conceptual framework.

References

1. Asemi, A., Zahra, K., Hasan Ashrafi, R.: Using LibQUAL+TM to improve services to libraries. A report on academic libraries of Iran experience. *The Electronic Library* 28(4), 568–579 (2010)
2. Junaida, I., Mahadir, L., Amin, S.H.M., Shari, S., Jusoff, K.: Influence of reliability dimension on service quality performance in northern region Malaysian University Academic Library. *Canada Social Science* 5(4), 113–119 (2009)
3. Arshad, A., Amen, K.: Service quality of the University of the Punjab's libraries. An exploration of users' perceptions. *Performance Measurement and Metrics* 11(3), 313–325 (2010)
4. Bertot, J.C.: Libraries and Networked Information Services: Issues and Consideration in Measurement. *Performance Measurement and Metrics* 5(1), 11–19 (2004)
5. Poll, R., te Boekhorst, P.: *Measuring Quality: Performance Measurement in Libraries*, 2nd rev. edn. IFLA Publications 127, Saur, München (2008)
6. Grove, S.J., Fisk, R.P., John, J.: The futures of services marketing: Forecasts from ten services experts. *Journal of Services Marketing* 17(2), 107–121 (2003)
7. Mathur, M.: Advent of Digital Libraries and measuring their performance: a review. *DESIDOC Bulletin of Information Technology* 25(2), 19–25 (2005)
8. Poll, R.: Quality measure for special libraries. In: *World Library and Information Congress: 73rd IFLA General Conference and Council, Durban (2007)*, <http://www.ifla.org/IV/ifla73/papers/152-Poll-en.pdf>
9. Hsieh, L.-F., Chin, J.B., Wu, M.C.: The performance indicators of university e-library in Taiwan. *The Electronic Library* 22(4), 325–330 (2004)
10. Shoeb, Z.H.: Identifying service superiority, zone of tolerance and underlying dimensions. Service quality attributes in a private university library in Bangladesh. *Library Review* 60(4), 293–311 (2011)
11. Reza, J., Hossein Sayyadi, T.: Priqritizing academic library service quality indicators using fuzzy approach. Case study: libraries of Ferdowsi University. *Library Management* 30(4/5), 319–333 (2009)
12. Kaur, K.: Service quality and customer satisfaction in academic libraries. Perspectives from a Malaysian University. *Library Review* 59(4), 261–273 (2010)
13. Thompson, B., Kyrrillidou, M., Cook, C.: Item sampling in service quality assessment surveys to improve response rates and reduce respondent burden. The “LibQUAL + Lite” example. *Performance Measurement and Metrics* 10(1), 6–16 (2009)
14. Balague, N., Saarti, J.: Bechmarking quality systems in two European academic libraries. *Library Management* 30(4/5), 227–239 (2009)
15. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64(1), 12–40 (1988)
16. Nitecki, D.A.: Changing the concept and measure of service quality in academic libraries. *The Journal of Academic Librarianship* 22(3), 181–190 (1996)
17. Bertot, J.C.: Libraries and Networked Information Services: Issues and Consideration in Measurement. *Performance Measurement and Metrics* 5(1), 11–19 (2004)
18. Vassiliadis, B., Stefani, A., Tsaknakis, J., Tsakalidis, A.: From application service provision to service-oriented computing: A study of the IT outsourcing evolution. *Telematics and Informatics* 23(4), 271–293 (2006)

An Adaptive IEEE 802.15.4 Data Transmission Scheme for Smart Grid Applications

Invited Paper

Zhongzhi Che¹, Jianhua He², Zuoyin Tang², Yi Zhou^{3,*}, and Kai Chen⁴

¹ Yantai Institute of China Agricultural University, Yantai China

² School of Engineering and Applied Science, Aston University, UK

³ College of Engineering, Swansea University, Swansea, UK

zy_21th@yahoo.com.cn

⁴ School of Information Security & Engineering, Shanghai Jiaotong University, China

Abstract. IEEE 802.15.4 networks (also known as ZigBee networks) has the features of low data rate and low power consumption. In this paper we propose an adaptive data transmission scheme which is based on CSMA/CA access control scheme, for applications which may have heavy traffic loads such as smart grids. In the proposed scheme, the personal area network (PAN) coordinator will adaptively broadcast a frame length threshold, which is used by the sensors to make decision whether a data frame should be transmitted directly to the target destinations, or follow a short data request frame. If the data frame is long and prone to collision, use of a short data request frame can efficiently reduce the costs of the potential collision on the energy and bandwidth. Simulation results demonstrate the effectiveness of the proposed scheme with largely improve bandwidth and power efficiency.

1 Introduction

IEEE 802.15.4 standard has been a strong candidate technique for wireless sensor networks and can find many applications to smart grid. [1][2]. Two types of medium access control algorithms are specified in the standard: unslotted and slotted carrier sensing multiple access and collision avoidance (CSMA/CA) algorithm [1]. Slotted CSMA/CA is used in the beacon enabled mode and unslotted CSMA/CA is used in non beacon-enabled mode [1]. Beacon-enabled mode has better energy efficiency due to the periodic superframe structure with which sensor devices can switch off power during the inactive period. Beacon-enabled mode is also flexible and efficient on supporting real-time applications if compared to non beacon-enabled mode.

In this paper we propose an adaptive CSMA/CA access control based data transmission scheme, for applications which may have heavy traffic loads such as smart grids [3][4]. In the proposed scheme, the PAN coordinator adaptively broadcast a frame length threshold, which is used by the sensors to decide

* Corresponding author.

whether a data frame should be transmitted directly to the target destinations or should follow a short data request frame. If the data frame is long and is prone to collision, use of a short data request frame can efficiently reduce the costs of the potential collision on the energy and bandwidth.

2 General Transmission and Reception Procedure

With slotted CSMA/CA algorithm, data frames can only be transmitted from the coordinator to the devices (called downlink traffic) or from the devices to the coordinator (called uplink traffic). Fig.1 shows the general frame structure that is used by the IEEE 802.15.4 CSMA/CA MAC protocol [1]. Fig.2 shows more information on the frame control fields that are a part of the general frame structure [1]. The adaptive medium access control scheme is to be used only with the uplink traffic in the slotted CSMA/CA algorithm, and can be used for all the traffic in the unslotted CSMA/CA algorithm. In the unslotted CSMA/CA algorithm, traffic can be transmitted between the sensor devices in range, not necessarily to between only the devices and the coordinator. Next we will introduce the basic transmission procedure with slotted CSMA/CA algorithm.

Octets: 2	1	0/2	0/2/8	0/2	0/2/8	0/5/6/10/ 14	variable	2
Frame Control	Sequence Number	Destination PAN Identifier	Destination Address	Source PAN Identifier	Source Address	Auxiliary Security Header	Frame Payload	FCS
	Addressing fields							
MHR							MAC Payload	MFR

Fig. 1. General frame structure [1]

Bits: 0-2	3	4	5	6	7-9	10-11	12-13	14-15
Frame Type	Security Enabled	Frame Pending	Ack. Request	PAN ID Compression	Reserved	Dest. Addressing Mode	Frame Version	Source Addressing Mode

Fig. 2. Frame control fields [1]

2.1 Downlink and Uplink Traffic

The downlink traffic to a device will be indicated by the coordinator in the beacon frame by listing the addresses of the devices which have pending data [1]. A device which is aware of the pending data in the coordinator attempts to extract a data frame from the coordinator by sending a data request command with slotted CSMA/CA algorithm in the contention access period (CAP).

If the data request command is successfully by the coordinator, the coordinator sends back an acknowledgment frame, thus confirming its receipt. If the coordinator has enough time to determine whether the device has a frame pending before sending the acknowledgment frame, it sets the Frame Pending subfield

of the Frame Control field of the acknowledgment frame accordingly to indicate whether a frame is actually pending for the device. If this is not possible, the coordinator sets the Frame Pending subfield of the acknowledgment frame to one.

If a device has data to send to the coordinator, it will contend the channel access with the slotted CSMA/CA algorithm, which will be described in the next subsection. If the device successfully accesses the channel, it will send the data frame to the coordinator. On the successful receipt of the data frame, the coordinator will send an ACK frame to the device if an ACK is requested as indicated in the data frame.

2.2 Slotted CSMA/CA Algorithm

It is specified in the standard that all frames, except acknowledgment frames and any data frame that quickly follows the acknowledgment of a data request command, transmitted in the CAP shall use a slotted CSMA/CA mechanism to access the channel. In slotted CSMA/CA algorithm [1], each device maintains three variables for each transmission attempt: NB, W and CW. NB is backoff stage, representing the number of times the CSMA/CA algorithm is required to backoff while attempting the current transmission. W is backoff window, used to determine the number of backoff slots a device shall wait before assess a channel. CW is the contention window length, defining the number of backoff periods that need to be clear for CCA before a transmission. CW is initialized to 2 before each transmission and reset to 2 each time the channel is assessed to be busy.

Before each new transmission attempt, NB is initialized to 0 and W is initialized to W_0 . A backoff counter is set with value chosen as a random number of complete backoff periods in the range $[0, W_0 - 1]$. The backoff counter value decrements at every slot without sensing the channel. Once the backoff counter reaches 0, perform the first CCA (denote by CCA1) from the boundary of the backoff slots. If the channel is idle in CCA1, then decrement CW by one and perform the second CCA (denoted by CCA2) in the next backoff slot. If the channel is idle again during CCA2, transmit the frame on the boundary of the next backoff slot. If the channel is busy during either CCA1 or CCA2, reset CW to 2, increase NB by one and double W (ensuring W is no more than W_x). If NB is less than or equal to the allowed number of retries (denoted by m), repeat the above backoff and CCA processes. If NB is greater than m , terminate the CSMA/CA algorithm.

3 Adaptive Scheme with Slotted CSMA/CA Algorithm for Uplink Traffic

In the standard, the uplink data frames are required to use slotted CSMA/CA mechanism to access the channel. In the frame exchanges, the uplink data frame is transmitted by a tagged source device after the device win the competition on channel access with the slotted CSMA/CA algorithm. If the data frame does

not collide with frames from the other competing devices, the coordinator will successfully receive the uplink data frame and will send back an ACK in acknowledgment mode.

The idea of the proposed adaptive slotted CSMA/CA algorithm based data transmission scheme is described as below. After a tagged device attempting to transmit a uplink data frame succeeds on the CCA2, if the data frame is longer than a frame length threshold (a value determined by the coordinator), the tagged device transmits a uplink data request (UDR) frame to the coordinator, instead of transmitting the long data frame to the coordinator directly. Note that the UDR frame has the meaning only for slotted CSMA/CA algorithm. For the unslotted slotted CSMA/CA algorithm, there is no direction for traffic and data request frame will be used instead of UDR frame. If the UDR frame is successful, the coordinator sends back an ACK confirming the receipt of the UDR frame. Then the tagged device can transmit the long data frame to the coordinator, and the data frame is ensured to experience no collision with other frames if there is no hidden terminal problem. The collision will only happens on the UDR frame (which is very short) instead of the long data frame. In this way, it possible to avoid the big waste of energy and bandwidth due to the collision of the long data frame, therefore efficiently save energy and bandwidth.

The proposed adaptive data transmission scheme can be illustrated in Fig.3 for non-acknowledgment mode and Fig.4 for acknowledgment mode. The coordinator will adaptively determine the frame length threshold, which will be used by the devices to determine whether a uplink data frame should be transmitted directly to the coordination or following a UDR frame. The proposed adaptive slotted CSMA/CA scheme can be implemented with full compatibility to legacy implementation of the slotted CSMA/CA algorithm based on the frame structures specified in the standard.

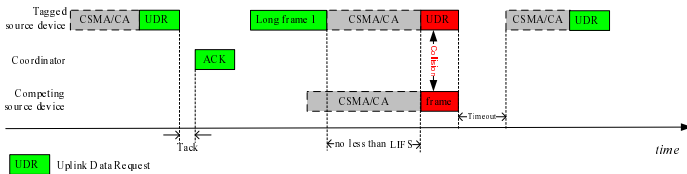


Fig. 3. Frame exchanges of proposed scheme with non-ACK mode

4 Performance Evaluation

We assume a one-hop WPAN, where one PAN coordinator communicates with some sensor devices, as shown in Fig.5. For simplicity, we assume that all the devices associated with the coordinator have only uplink traffic. The channel access parameters are set with $W_0=2^3$, $W_x=2^5$ and $m=4$. The PHY data rate is 250 kbps.

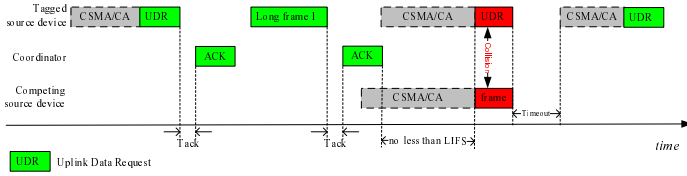


Fig. 4. Frame exchanges of proposed with ACK mode

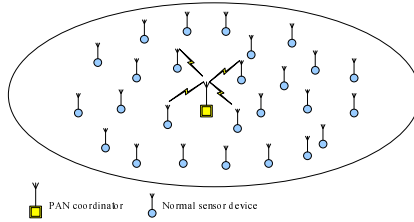


Fig. 5. An example network scenario

The performances of normalized throughput, energy consumption and frame success probability are plotted versus the number of devices in Fig.6, Fig.7 and Fig.8, respectively. In the figures, label “Std” represents the results obtained with the standard slotted CSMA/CA algorithm. Label “New” represents the results obtained with the proposed adaptive scheme, in which each data frame is simply required to follow a UDR. The data payloads of $L_d=20$ bytes and $L_d=100$ bytes are investigated. It is noted that the maximal data payload in a MAC frame is 127 bytes.

From the presented figures, we can observe that when the length of data frame is large (e.g., $L_d=100$ bytes), the new algorithm can achieve much better performances than the standard algorithm with the increased number of competing devices (e.g., 20 devices). For example, for $L_d=100$ bytes, 20 devices, the

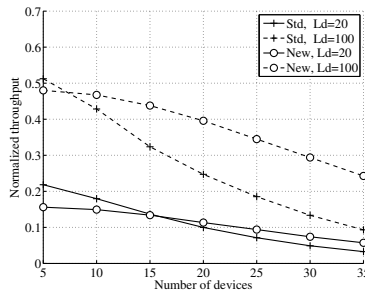


Fig. 6. Throughput with non-acknowledgment mode, $L_d=20$ bytes (solid line) and $L_d=100$ bytes (dashed line)

throughput of the standard algorithm and the new algorithm is 0.25 and 0.4 in Fig. 6 respectively. The improvements on throughput and energy consumption are significant, and show the effectiveness of the proposed algorithm.

For the short frame length (e.g., data payload $L_d=20$ bytes), there is no obvious improvement by using UDR. Therefore the coordinator can adaptively set the frame length threshold for the use of UDR, to improve the performances of the slotted CSMA/CA algorithm. Simulation results demonstrated the effectiveness of the proposed scheme.

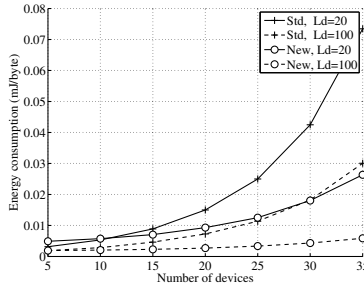


Fig. 7. Energy with non-acknowledgment mode, $L_d=20$ bytes (solid line) and $L_d=100$ bytes (dashed line)

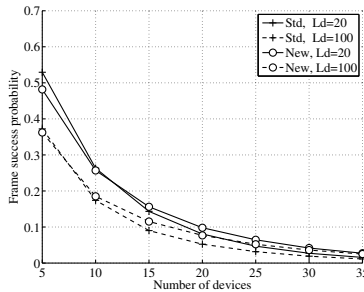


Fig. 8. Reliability with non-acknowledgment mode, $L_d=20$ bytes (solid line) and $L_d=100$ bytes (dashed line)

Acknowledgement. The work was partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) with grant reference number EP/1010157/1.

References

1. IEEE Std 802.15.4, Wireless medium access control and physical layer specifications for low-rate wireless personal area networks (2006)
2. Lu, G., et al.: Performance evaluation of the IEEE 802.15.4 MAC for low-rate low-power wireless networks. In: Workshop EWCN 2004 (April 2004)

3. Gungor, V., Lu, B., Hancke, G.: Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics* 57(10), 3557–3564 (2010)
4. Farhangi, H.: The path of the smart grid. *IEEE Power and Energy Magazine* 8(1), 18–28 (2010)
5. He, J., Tang, Z., Chen, H., Wang, S.: An accurate markov model for slotted csma/ca algorithm in IEEE 802.15. 4 networks. *IEEE Communications Letters* 12(6), 420–422 (2008)
6. He, J., Tang, Z., Chen, H., Zhang, Q.: An accurate and scalable analytical model for IEEE 802.15. 4 slotted CSMA/CA networks. *IEEE Transactions on Wireless Communications* 8(1), 440–448 (2009)

Modeling Impact of Both Frame Collisions and Frame Corruptions on IEEE 802.15.4 Channel Access for Smart Grid Applications

Zhongzhi Che¹, Jianhua He², Yi Zhou^{3,*}, Zuoyin Tang², and Chao Ma³

¹ Yantai Institute of China Agricultural University, Yantai China

² School of Engineering and Applied Science, Aston University, UK

³ College of Engineering, Swansea University, Swansea, UK
zy_21th@yahoo.com.cn

Abstract. IEEE 802.15.4 standard has been recently developed for low power wireless personal area networks. It can find many applications for smart grid, such as data collection, monitoring and control functions. The performance of 802.15.4 networks has been widely studied in the literature. However the main focus has been on the modeling throughput performance with frame collisions. In this paper we propose an analytic model which can model the impact of frame collisions as well as frame corruptions due to channel bit errors. With this model the frame length can be carefully selected to improve system performance. The analytic model can also be used to study the 802.15.4 networks with interference from other co-located networks, such as IEEE 802.11 and Bluetooth networks.

1 Introduction

IEEE 802.15.4 has been recently standardized for low rate wireless personal area networks (LR-WPANs) [1]. Compared to Bluetooth 802.15.4 networks has much lower power consumption, lower cost and more flexible networking. It has been a strong candidate technique for wireless sensor networks and can find many applications to smart grid [2][3]. Smart grid is an upgrade of 20th century power grid. It is capable of delivering electricity in more optimal ways from suppliers to consumers using two-way digital technology to control appliances at consumers' homes to save energy, reduce cost and increase reliability and transparency. It has been promoted as a way of addressing energy independence, global warming and emergency resilience issues.

For the smart grid applications a key component is the intelligent monitoring system, which can provide up to date grid and energy usage statuses. IEEE 802.15.4 sensors could be used to do the monitoring tasks. As the 802.15.4 sensors may locate at different distances from the network coordinators, the frames sent from the sensors may be subject to both frame corruptions due to weak receive power and frame collisions due to channel access. Therefore it is important to

* Corresponding author.

model the impact of both frame corruptions and collisions on 802.15.4 channel access for smart grid applications. In the literature many analytic models have been proposed to capture the throughput and energy consumption performances of single IEEE 802.15.4 network with either saturated or unsaturated traffic. The authors proposed an accurate Markov mode in [4][5]. However, to the best of our knowledge our research on modeling the impacts of both frame collision and frame corruption on 802.15.4 network performance is the first of its kind reported in the literature. With our analytic model the frame length can be carefully selected to improve system performance.

2 Channel Access Schemes

802.15.4 slotted CSMA/CA algorithm operates in unit of backoff slot. One back-off slot is 20 symbols. In the non-ACK mode every node in the network maintains three variables for each transmission attempt: NB , W and CW . NB is the back-off stage, representing the backoff times that each node need to while it trying to transmit data in each transmission. W is the backoff window, representing the number of backoff slots that a node need to backoff before CCA. CW is the contention window length, used to determine how many backoff periods for the CCA before transmission. CW will be set to two before each transmission and reset to two when the channel is sensed busy.

Before each new transmission attempt, NB will set to zero and W will set to W_0 . The backoff counter will chose random number form $[0, W_0 - 1]$ and it decrements every backoff slot without sensing channel until reaches zero. The first CCA (denoted by CCA1) will preformed when backoff counter reaches zero. If channel is idle at CCA1, CW decrements by one and the second CCA (denoted by CCA2) will performed after CCA1. If channel is idle for both CCA1 and CCA2, the frame will be transmitted in the next slot. If either CCA1 or CCA2 is busy, CW will rest to two, NB increases by one, and W will be doubled but not exceed W_x . If NB is less than or equal to an allowed number of retries (denoted by m), the above backoff and CCA processes are repeated. If NB is greater than m , CSMA/CA algorithm ends.

3 System Model

We assume there are N devices in the networks and each node has saturated traffic to send to its coordinator with non-ACK mode. Each data frame has a fixed length which requires L slots for transmission over the channel. Let the length of data payload in a data frame denoted by L_d slots. We assume that the channel is ideal which means a data frame could be correctly received if it is not collided. The superframe is assumed to consist only contention access period.

3.1 Frame Corruption Probability

The physical layer of IEEE 802.15.4 standard at 2.4 GHz uses offset quadrature phase shift keying modulation [1]. Let P_{rx} , P_{no} and P_{int} be the signal power,

the noise power, and interference power at the 802.15.4 receiver. Then the signal to interference and noise ratio (SINR) and bit error rate (BER) (denoted by p_b) of 802.15.4 node can be calculated by [6]:

$$SINR = 10 \log_{10} \frac{P_{rx}}{P_{no} + P_{int}} + P_{gain} \quad (1)$$

and

$$p_b = Q(\sqrt{2\alpha SINR}). \quad (2)$$

where P_{gain} is processing gain, and $\alpha = 0.85$. From the bit error rate p_b and frame length frame corruption probability (denoted by p_{corr}) can be calculated.

3.2 Frame Collision Probability

According to the idea of performance modeling in [4] the overall channel states sensed by each node can be modeled by a renewal process, which starts with an idle period and followed by a fixed length of L slots. The idle period depends on the random backoff slots and the transmission activities from each node. The slotted CSMA/CA operations of each individual node could be modeled by an Markov chain with finite states. Let p_k denote the probability of a transmission from devices in the network other than a tagged device in the network starting after exactly k^{th} idle slots since the last transmission, where $k \in [0, W_x + 1]$ [4]. The transmission probability of a node in a general backoff slot in the renewal process can be calculated with the Markov chain constructed for each node.

Without loss of generality we consider a tagged node in the networks. For the tagged node its Markov chain consists of a number of finite states, each corresponding to a state of the CSMA/CA state in one slot. These finite states are introduced below. Let \bar{M} denote the steady state probability of a general state M in the Markov state space.

Busy State. Denoted by $B_{i,j,l}$, during which at least one node other than the tagged node transmits the l th part of a frame of L slots, with the backoff stage and backoff counter of the tagged node being i and j , respectively, where $i \in [0, m]$, $j \in [0, W_i - 1]$, and $l \in [2, L]$, W_i is the minimum of $2^i W_0$ and W_x [4].

$$\begin{aligned} \bar{B}_{0,j,2} &= \sum_{k=2}^{W_0-1} p_k \bar{K}_{0,j+1,k} + \frac{1}{W_0} \sum_{k=2}^{W_m} p_k (\bar{K}_{m,0,k} + \bar{C}_{m,k}) \quad i=0, j \in [0, W_0 - 1] \\ \bar{B}_{i,j,2} &= \sum_{k=2}^{W_i-1} p_k \bar{K}_{i,j+1,k} + \frac{1}{W_i} \sum_{k=2}^{W_i-1} p_k (\bar{K}_{i-1,0,k} + \bar{C}_{i-1,k}) \quad i \in [1, m], j \in [0, W_i - 1] \\ \bar{B}_{i,j,l} &= \begin{cases} \bar{B}_{0,j+1,l-1} + \bar{B}_{m,0,l-1}/W_0, & i=0, j \in [0, W_i - 1] \\ \bar{B}_{i,j+1,l-1} + \bar{B}_{i-1,0,l-1}/W_i, & i \in [1, m], j \in [0, W_i - 1] \end{cases} \end{aligned} \quad (3)$$

where states $K_{i,j,k}$ are backoff states to be introduced next.

Backoff State. Denoted by $K_{i,j,k}$, during which the tagged node backoff with backoff counter being j at backoff stage i , after k idle slots since the last transmission, where $i \in [0, m]$, $j \in [0, W_i - 1]$, and $k \in [0, W_i - 1]$ [4].

$$\begin{aligned} \bar{K}_{0,j,0} &= \bar{B}_{0,j+1,L} + (\bar{B}_{m,0,L} + \bar{T}_L)/W_0 & i = 0, j \in [0, W_0 - 1] \\ \bar{K}_{i,j,0} &= \bar{B}_{i,j+1,L} + \bar{B}_{i-1,0,L}/W_i & i \in [1, m], j \in [0, W_i - 1] \\ \bar{K}_{i,j,k} &= \begin{cases} \bar{K}_{i,j+1,k-1}, & k \in [1, 2] \\ (1 - p_{k-1})\bar{K}_{i,j+1,k-1}, & 3 \leq k \leq W_i - 1 \end{cases} \end{aligned} \quad (4)$$

Sensing State. Denoted by $C_{i,k}$, during which the tagged node performs CCA2 at the i th backoff stage, after k idle slots since the last transmission, where $i \in [0, m]$ and $k \in [1, W_i]$ [4].

$$\bar{C}_{i,k} = \begin{cases} \bar{K}_{i,0,k-1}, & k \in [1, 2] \\ (1 - p_{k-1})\bar{K}_{i,0,k-1}, & k \in [3, W_i] \end{cases} \quad (5)$$

Initial Transmission State. Denoted by $X_{i,k}$, during which the tagged node starts to transmit a frame at backoff stage $i \in [0, m]$, after $k \in [2, W_i + 1]$ idle slots since the last transmission [4].

$$\bar{X}_{i,k} = \begin{cases} \bar{C}_{i,k-1}, & k = 2 \\ (1 - p_{k-1})\bar{C}_{i,k-1}, & k \in [3, W_i + 1] \end{cases} \quad (6)$$

Transmission State. Denoted by T_l , during which the tagged node transmits the l th part of a frame, where $l \in [2, L]$. The first part is transmitted in the state $X_{i,k}$ [4].

$$\bar{T}_l = \begin{cases} \sum_{i=0}^m \sum_{k=2}^{W_i+1} \bar{X}_{i,k}, & l = 2 \\ \bar{T}_{l-1}, & l \in [3, L] \end{cases} \quad (7)$$

The transmission probability τ_k that the tagged node transmits after exactly k idle slots since the last transmission in channel can be computed by $\tau_k = 0$, for $k \in [0, 1]$, and for $k \in [2, W_x + 1]$ [4].

$$\tau_k = \sum_{i=0}^m \bar{X}_{i,k} / \sum_{i=0}^m [\bar{X}_{i,k} + \bar{C}_{i,k} + \sum_{j=0}^{W_i-1} \bar{K}_{i,j,k}] \quad (8)$$

With the above expressions derived for transmission probability τ_k , we can calculate channel busy probability p_k for $k \in [0, W_x + 1]$. We have

$$p_k = 1 - (1 - \tau_k)^{N-1}. \quad (9)$$

Since the balance equations for all steady state probabilities and expressions for $p_k, k \in [0, W_x + 1]$ have been derived, the Markov chain for the tagged node can be numerically solved. After the Markov chains are solved, we can calculate the throughput of individual network and the overall system. We have

$$S = NL_d \sum_{i=0}^m \sum_{k=1}^{W_i} C_{i,k-1} (1 - p_{k-1})(1 - p_k)(1 - p_{corr}). \tag{10}$$

Note that impact of both frame collision and corruption have been modeled in the calculation of the system throughput.

4 Numeric Results

A discrete event simulator is implemented for IEEE 802.15.4 CSMA/CA algorithm to verify the proposed analytical model. The MAC parameters are configured with default values: $W_0 = 2^3, W_x = 2^5,$ and $m = 4.$ We assume that the overhead of the header L_h in a frame is 1.5 slots in the simulation, and the data length with MAC and PHY layer header is $L = L_d + L_h.$ Each simulation result was obtained from the average of 15 simulations and transmitting 10^5 frames.

The performances of normalized throughput is plotted versus the number of devices in Fig.1, and Fig.2 for data frame length 60 bytes and 100 bytes, respectively. It is noted that the maximal data payload in a MAC frame is 127 bytes. In the figures, label ‘‘Sim’’ represents the simulation results for the slotted CSMA/CA algorithm. In both figures, results with and without frame corruption are presented. For the frame corruption case the SINR is set to 5 dB.

From the presented figures, it is observed that the analytic results agree well with simulation results. We can also observe that when there is no frame corruption (which means the SINR is sufficiently high), the longer the data frame, the

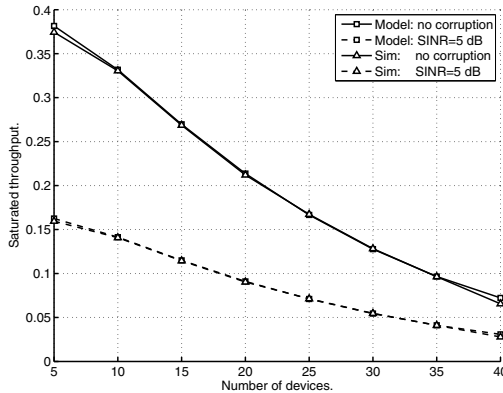


Fig. 1. Throughput with non-acknowledgment mode, $L_d=60$ bytes (solid line) and $L_d=100$ bytes (dashed line)

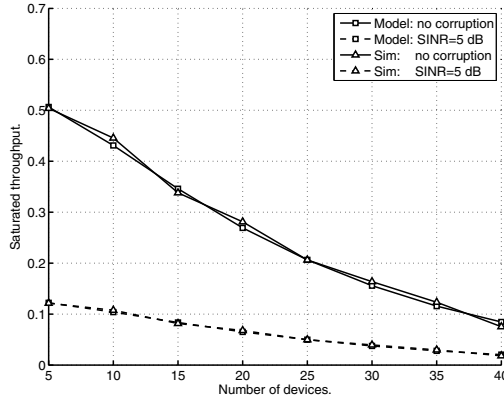


Fig. 2. Energy with non-acknowledgment mode, $L_d=60$ bytes (solid line) and $L_d=100$ bytes (dashed line)

higher throughput efficiency for the channel access scheme. This is mainly because of the fixed physical and MAC layer overhead. However, when there are frame corruptions, the longer the data frame, the more likely the frame to be corrupted. It is observed that when the SINR is low throughput efficiency with shorter frame length is higher than that with longer frame length. Therefore when both frame collision and frame corruption are present in the 802.15.4 networks, a careful selection of frame length may help improve the system performance.

Acknowledgement. The work was partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) with grant reference number *EP/1010157/1*.

References

1. IEEE Std 802.15.4, Wireless medium access control and physical layer specifications for low-rate wireless personal area networks (2006)
2. Gungor, V., Lu, B., Hancke, G.: Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Trans. Industrial Electronics* 57(10) (2010)
3. Farhangi, H.: The path of the smart grid. *IEEE Power and Energy Magazine* 8(1), 18–28 (2010)
4. He, J., Tang, Z., Chen, H., Wang, S.: An accurate markov model for slotted csma/ca algorithm in IEEE 802.15.4 networks. *IEEE Communications Letters* 12(6), 420–422 (2008)
5. He, J., Tang, Z., Chen, H., Zhang, Q.: An accurate and scalable analytical model for IEEE 802.15.4 slotted CSMA/CA networks. *IEEE Transactions on Wireless Communications* 8(1), 440–448 (2009)
6. Shin, S., Park, H., Choi, S., Kwon, W.: Packet error rate analysis of ZigBee under WLAN and Bluetooth interferences. *IEEE Transactions on Wireless Communications* 6(8) (August 2007)

Training Schedule Management Application Using Various Constraints

Orawan Mantrawong and Chantana Chantrapornchai

Department of Computing, Faculty of Science, Silpakorn University, Thailand 73000
ctana@su.ac.th

Abstract. In this paper, we develop an application for training schedule management. The tool suggests the training schedule for a particular trainee. There exists a database which contains the data about the trainee's skills and courses. The tool uses the constraint selection to arrange the schedule. The training schedule is then reported and suggested to the training company.

Keywords: Training schedule algorithm, Scheduling application, Training management.

1 Introduction

Training course arrangement is important to the success of a company. For large companies where there are many employees, each employee may have different skills at different levels. The training is needed for each employee to improve his knowledge so that he can work effectively. Since each training consumes time and have expenses, the proper training courses need to be selected so that the company can save time and cost to train the employee.

In this work, we are interested in training in creating such a training schedule. Given that the company has a database of employees as well as their skills, the training company offers a set of courses of various levels and requires different backgrounds, the schedule is customized to an employee. Such a problem is known to be a typical optimization problem with various constraints. We develop a heuristic to select courses and construct a sequence of course to be taken for each employee to minimize the time spent.

We also develop a program which contains the user interface to input the employee names, skills, as well as the courses. The tool generates reports to suggest the training schedule. The test data for the prototype is assumed and the course samples are obtained from one of public organizations.

It is known that the general scheduling problem is NP-complete. There have been many heuristics for solving such a problem such as using constraints. Aickelin, U. and Dowsland, K. used genetic algorithms for matching a mall layout and tenants [1]. Burke, E. used Tabu search scheduling nurse time table [2]. Chen, J.G.,and Yeung, T.W., presented an expert system for nurse scheduling [4]. Many works studied scheduling approaches in many applications. For example, the work of Wanichapong et.al.[10] uses constrained-based for finding bus driver and bus crew schedules. The

constraints are divided into hard and soft constraints. The method considered the hard constraint, then the soft constraints, the conflicting constraints and weight values compared to the goal. The work by Tangtriratanakul [9] consider the project schedule by using a mixed of ant colony and genetic algorithms. The work by Rattanadamrongagsorn [8] presented the construction scheduling using the constraint satisfaction problem. Cheang, B. et.al. [3] presented the approach to create a schedule for nurses in a hospital. Lin Wu et.al. [7] uses fuzzy set theory to develop a training schedule that considers emotions and ability of a trainee to create a flexible training schedule time. Chiou et.al. [6] used a genetic algorithm and matrices to develop training courses. However, he did not consider the attributes of the curriculum.

In this paper, like the work in [6] but we focus the application on the training schedule. Though there are many commercial project management tools, we would like to develop one that is customized for a specific company. The tool we develop will be useful to create a set of training courses for our personnel as well as developing a curriculum in the institution. Also, it stores a database of personnel skill and course curriculum. The application is a prototype which will be extended to perform various scheduling and assignment approaches.

The paper is organized as follows. Next section, we introduce some background on scheduling and the problem definition. In Section 3, the heuristic of the training scheduling creation is presented. Section 4 presents the example. Finally Section 5 concludes the work.

2 Backgrounds

Regular project management problems require precedence between tasks and constraints in resource requirement [8]. The goal of the schedule is to minimize the total time of the schedule and create a proper legal sequence of tasks.

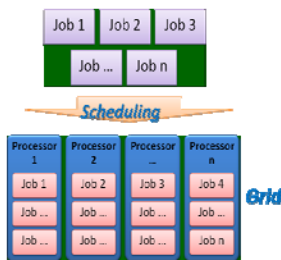


Fig. 1. Example of scheduling model for processors [8]

Figure 1 shows an example of traditional scheduling on multiprocessors. Normally to create a schedule, the direct acyclic graph is constructed to show the precedence of tasks. Each node corresponds to a task. Each node has an attributed as computation which may not be the same. Then the tasks are allocated to a processor according to the precedence order. The goal is to minimize the total time of the schedule. Figure 2 shows an example of the approach.

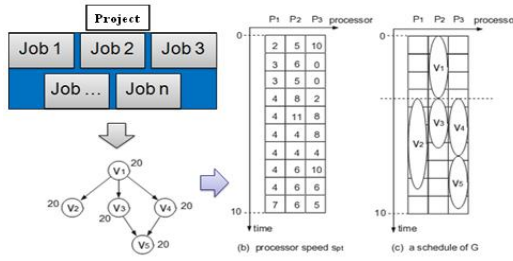


Fig. 2. Schedule and its direct acyclic graph [8]

In our paper, we focus on the course schedule problem. The problem is described as follows. Given a set of employees (v_i), each of which as a skill a_{ij} , for $i, j \geq 1$, a set of courses (c_i), a set of relation between c_i, c_j , for any i, j , implying that the c_i has to be taken before c_j , a set of required skills (s_{ij}) by course c_i , and a list of courses (x_i) to needed to be studied by the employee, we would like to find a sequence of courses $\{x_i\}$ that are suggested to the employees to be trained.

Assume for flexibility that the skills are only suggested by the course instructors. It is a soft constraint that the employee may not take the course even though they do not have all skills. When the employee finishes the course, he obtained the missing skills as well. Also, the employee has a preference list of courses ($rank(x_i)$) to be taken. This list will be used to solve the tie when there are more than one course to be suggested next.

3 Data Structures and Heuristics

In the software, we store data for each course, and employee as in the E-R diagram.

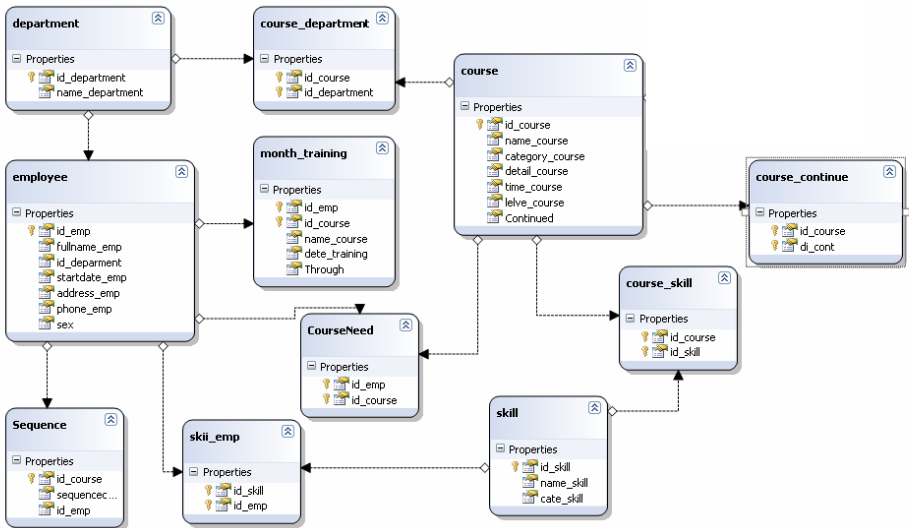


Fig. 3. Data representation

Employee tables stores employee information. It is associated with CourseNeed and Skill_Emp tables. Course table has information about the course. The important property of the course is the time for each course, the level of the course. It is associated with Course_Skill and Course_Continue tables. Course_Continue table states whether the course has the next course, for example, Programming I, Programming II etc. Sequence table is the schedule table.

To search for proper courses, our heuristic orders the constraints to be checked as following.

1. Matching phase: Employ skill a_{ij} is checked against the skill of each course s_{ij} for every employee and course. The selected course is the course that matches most of the skill of the employee. The ideal situation is all the skills match. This step sort the selected courses by the number of matched skills.
2. Prerequisite phase: In case that the selected course requires the prerequisite course(s) and the employee has not passed them all. The next course in the sorted list is considered. Also, the level or continue course is checked here. The course with level 0 should be taken before level-1 course etc.
3. Order phase: In case there are still more than one courses obtained by 2), the employee may have a list of preferences of courses to be taken. Therefore, this list is considered in this step.
4. If there is a continue course, the continue course is also suggested in the next round.

After the course is assumed to be selected to train, there will be new skills obtained. Then the step 1 is repeated again until there are no new courses to be derived.

4 Examples

Assume that an employee has the following skill. The highlight is the name of the employee. Underneath, there is a list of the employee skills.

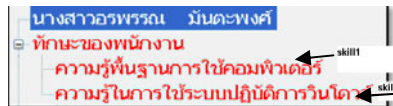


Fig. 4. Employee skills

รหัสวิชา	course name	preferenc
3	31-217: Intermediate Excel 2007 for Special Function and Database	1
4	31-218: Intermediate Excel 2007 for PivotTable and PivotChart (Level II)	2
6	31-308 Advanced Excel 2007 for ADO Programming and Workshop VBA	3
7	31-307 Advanced Word 2007 for Professional Document and Macro	4
5	31-304: Advanced Excel 2007 for Design Macro and VBA Programming (Level I)	5

Fig. 5. Sample employ courses' list

Figure 8 shows the example of prerequisites from the data set in one public organization. The edge shows course prerequisite and the attached rectangle shows the skilled needed.

In step 3, the preference in column “preference” in Figure 4 is used.

Consider the example in Figure 8 and Figure 9 showing a list of courses and preferences.

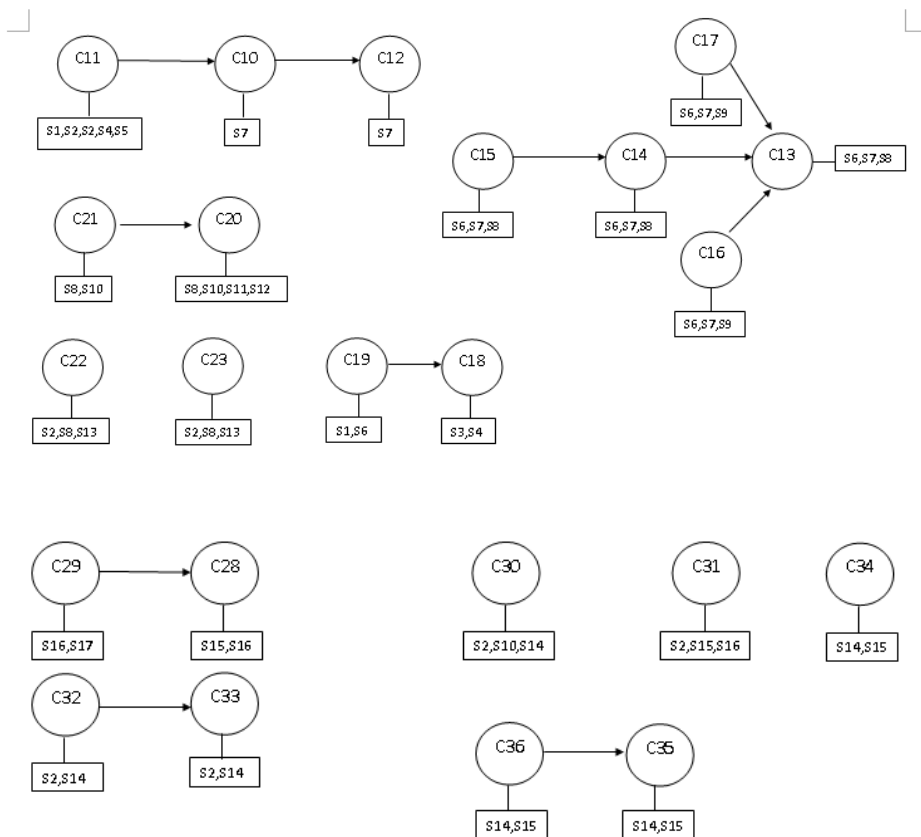


Fig. 8. Example of course prerequisites

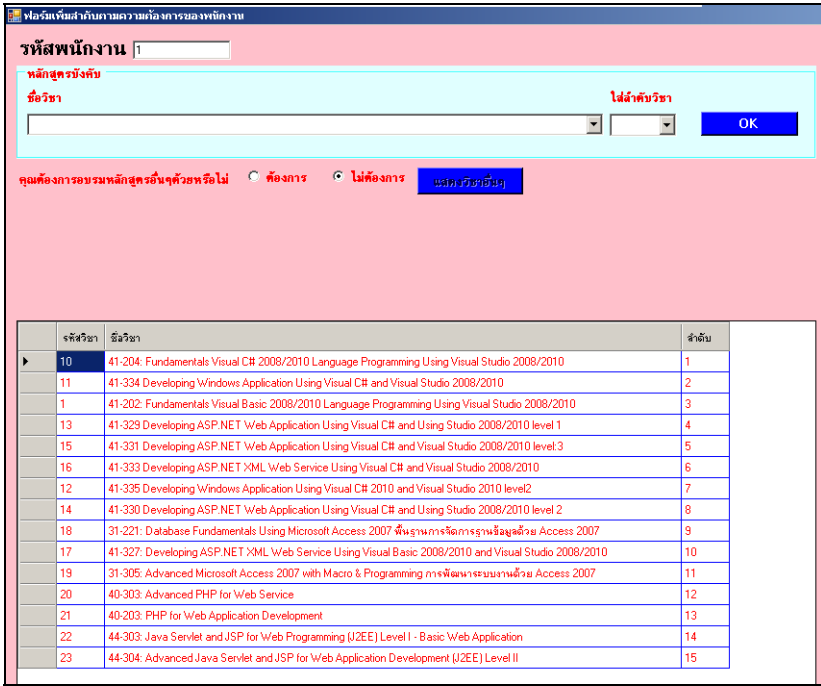


Fig. 9. Example of course needed

The output of the software is the sequence of courses to be trained in Figure 10(b). Figure 10(a) shows the prerequisites of some courses in this list.

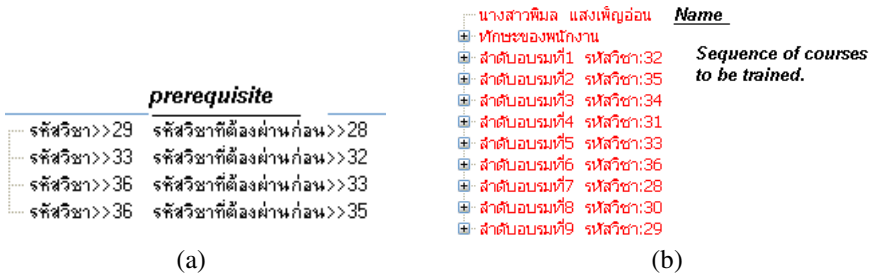


Fig. 10. Reports. (a) Prerequisites (b) Course sequences

5 Conclusion

The work presents an application for training schedule management. It includes a heuristic to create a sequence of training courses. The database of employ skill and courses are constructed. The employee inputs the courses he is interested in. Also, he gives the preferences of the courses. The curriculum of the available courses are

assumed. Each course has prerequisites and maybe the level of the courses. It also includes the skills needed by the course. The heuristic orders the courses by the prerequisite, level, skill, and preferences. The heuristic is integrated to a software to be used by a training department.

References

- [1] Aickelin, U., Dowsland, K.: Enhanced Direct and Indirect Genetic Algorithm Approaches for a Mall Layout and Tenant Selection Problem. *Journal of Heuristics* 8(5) (September 2002), doi:10.1023/A:1016536623961
- [2] Burke, E.K.: Scheduling Nurses Using a Tabu-Search Hyperheuristic. In: *Proc. of the 1st MISTA (2003)*, <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.3421>
- [3] Cheang, B., Li, H., Lim, A., Rodrigues, B.: Nurse rostering problems - a bibliographic survey. *European Journal Operation Research* (151), 447–460 (2003)
- [4] Chen, J.G., Yeung, T.W.: Hybrid expert-system approach to nurse scheduling. *Computers in Nursing* 11(4), 183–190 (1993)
- [5] Ferland, J.A., Berrada, I., et al.: Generalized Assignment Type Goal Programming Problem: Application to Nurse Scheduling. *Journal of Heuristics* 7(4) (July 2001), doi:10.1023/A:1011392328632
- [6] Juang, Y.S., et al.: An Adaptive Scheduling System with Genetic Algorithms for arranging employee training programs. *Expert Systems with Applications* 33, 642–651 (2007)
- [7] Lin, S.-S., Wu, M.-Y., Perng, C., Tsai, J.-T.: Application of fuzzy set theory on course selection in employee training program. In: *Proceedings of the Sixth Conference on Fuzzy Theory and its Application, Taichung, Taiwan (1998)*
- [8] Rattanadamrongagsorn, T.: Construction Scheduling using Constraint Satisfaction Paradigm. Individual Study, KMUTT (2003) (in Thai)
- [9] Tangtriratanakul, J.: Simulation Project Scheduling with Ant Colony Optimization & Genetic Algorithm. Individual Study, KU (2007) (in Thai)
- [10] Wanichapong, K., et al.: Solving bus crew scheduling problem: a constraint-based approach. Research Report (2008) (in Thai), <http://bal.buu.ac.th/bal2010/RR-2008-02>

Data Mining and Web-Based Children Shoe Suggestion System

P. Sudta, K. Kanchan, and Chantana Chantrapornchai

Department of Computing, Faculty of Science, Silpakorn University, Thailand, 73000
ctana@su.ac.th

Abstract. In this paper, we present a shoe suggestion system for children. The system stores the database about children's shoes from K1-K3 and P1-P3. The attributes considered are types of shoes, shoe sizes, brands. Given a type of children foot and age, the system suggested the type of shoes, brands, and sizes with the confidence level. The system uses the data mining technique for the classification and prediction.

Keywords: Shoe suggestion, data mining, decision tree, KNN, Neural net.

1 Introduction

For children, there are various kinds of attributes that are needed to consider for buying their shoes. The children shoes need to be comfort. The size needs to be suitable to them as well as the shape of the shoes. In this research, we are interested in shoe sizing and major attributes that affect the decision on buying shoes for Thai students. In particular, we would like to develop a prototype system that suggests proper shoes for small kids based on certain criteria.

Data mining is a popular technique used in classification and predictions. Many research work used it as a tool for decision support system. Palaniappan S. and Awang R. [7] developed a heart disease decision support system. It used naïve bayes, decision tree and neural net to model the classifications. It is developed as a web applications and used DMX query tool. C.Y. Ma, Frances, V Buontempo and X. Z Wang [6] used inductive data mining for historic data analysis. M.J. Aitkenhead [1] developed a co-evolution decision tree method. It combines an evolution method with the decision tree model for a better classification. Hudson S., and Ritchie B. [4] used cluster analysis for tourism in Alberta. They applied the model CRISP-DM. Jiao J. R et. al. [5] developed Kensei Mining which uses association rules to design user interface. Hsieh C., Huang S. [3] applied data mining technique to design new products. Apriori algorithm was used here. The attributes investigated are customer needs and product properties. Prassas G., et. al.[8] data mining for online shopping suggestion. Chen Y.L., Chen J.M., Tung C.W. [2] applied data mining to see the effect of shelf-space adjacency.

In our paper, we develop a children shoe suggestion system. The data mining engine is used. Particularly the decision tree technique is applied. We collect samples data from the neighborhood school from K1-K3 and P1-P3, each of which is 20-30

students. About the students, we collect the student age, shoe shape, sex. About their shoes, we collect the shoe sizes, brands, and types. Assume the types are student shoes, sport shoes and leisure shoes. We design the database to save all the student data and shoe attributes. After the modeling for classification is developed, it is integrated to the web-based system.

2 Backgrounds and Data Preparations

Data mining technique is used often with large database, dataware house etc. It is mainly applied in decision support systems for modeling and prediction. There are several kinds of data mining: classification, clustering, association, sequencing etc.

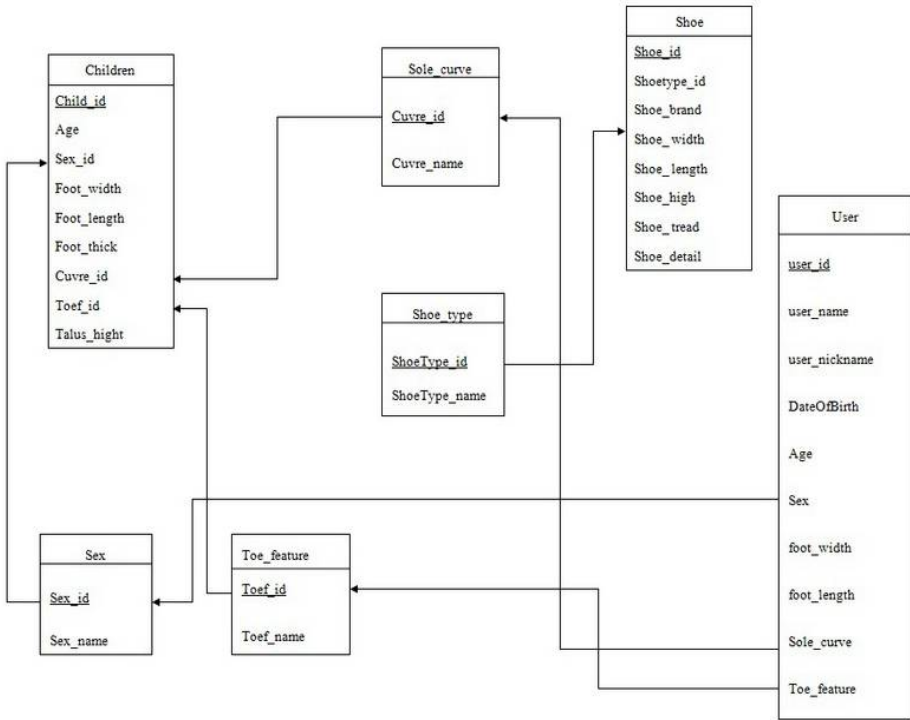


Fig. 1. E-R diagram of the collected data

Typically, for classification, the classifier model is needed. Data are divided into training set and test set. The training data is used to create the model. Then the test set applied for checking the model correctness. Until satisfied, the model is trained and adjusted by training data. Common techniques used in classification are decision tree, neural network, naïve bayes, etc.

In this work, we collect the data from one of the public kindergarden school. The school is a small size. There are 20 boys each in K1, K2 and K3 respectively. There are 30 boys in P1, P2, and P3 respectively. There are 20 girls each in K1, K2, and K3

respectively. Also, there are 30 girls each in P1,P2 and P3 respectively. There are 3 types of shoe types such as sport shoes, student shoes and leisure shoes. We collect 120 pairs of student shoes dividing into 60 boys' and 60 girl's, 70 pairs of sport shoes from 35 boys and 35 girls, 50 pairs of leisure shoes from 25 boys and 25 girls.

We design a form for collecting these data from students. The attributes that are specified are student information such as name, sex, age, foot information such as sole curve, toes characteristics, thickness, height. We also draw the foot figures for left and right sides for references. The shoe sizes for each type are collected as well. The sample form is shown in appendix. From the data collected, we design a E-R diagram which corresponds to the database table for the application in Figure 1. Then these data are divided into training set and test set as in Figure 2.

In Figure 2, the attributed used to train are age, sex, foot width, foot length, foot thickness, sold curve, foot type, toe type, and length upto the angle. The last two columns are the brands and shoe sizes.

3 Data Analysis

From the above data collection, we analyze the attributes of the feet as following.

-Foot length. The students for these ages mostly have the same foot length. There are very few samples that have a distinct foot size.

-Sole curve. Usually, there are three types of sole curve. However, since these are very small children, the sole curves are in only flat level and small curve.

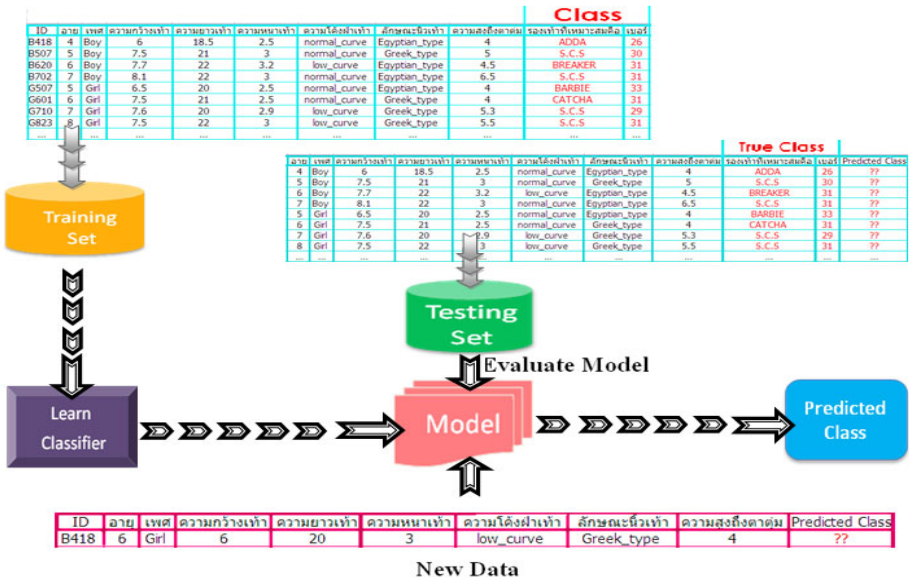


Fig. 2. Training the model

-Toe feature. Typically, there are three types: Egyptian style (where left one in the right side is the longest one and the others are shorter in the decreasing order), Greek style (where the second one is the longest) and the square style (where every finger is of the same height). From our data, most of the students are in the Greek style, then Egyptian style. Then the for the student shoe type, it can be matched with all kinds of feet. For the sport shoes, by inspecting the head of the shoes, it may be matched only Greek style feet. For the leisure shoes, by inspecting the head of the shoes, it may also be matched only Greek style feet.

From our data, we still have problems of insufficient data for children. Most likely, children in the same area and nationality, have the same style of feet. Also, the available shoes are very similar to each other. Then, the foot styles are not influence the classification so much.

4 Results

We use Weka (<http://www.cs.waikato.ac.nz/ml/weka/>) tool to create the decision tree. Figure 3 is an example result for student shoes. In this example, 1) foot length is the most important attribute. Then it is sole curve (small curve) and sex (female). We obtain CATCHA_31(6.0) which is the brand CATCHA, size 31, and there are 6 students on this leaf. In 2), also, foot length is the most important attribute. Then it is sole curve (small curve) and sex (male). The shoes is SCS brand whose size is 33 and there are 5 students in this leaf.

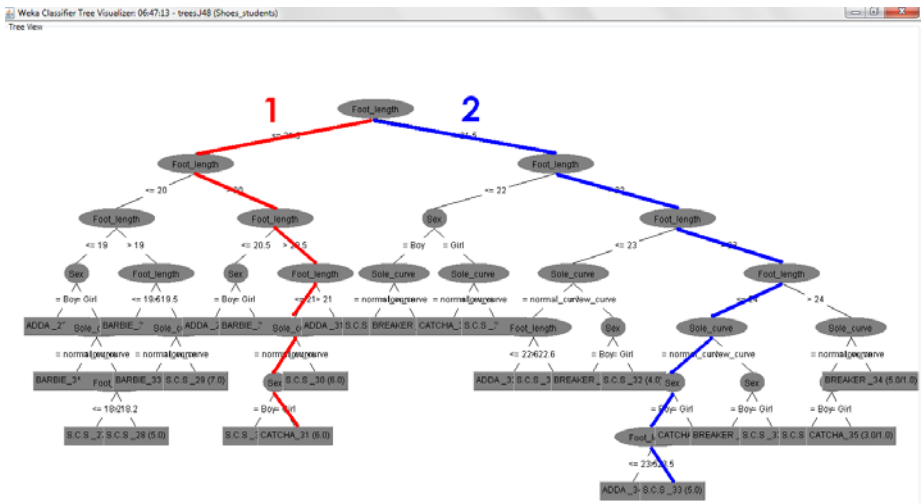


Fig. 3. Sample tree results for student shoe classification

In Figure 4, it is the sample results from sport shoes. In 1), only the foot length is used. In 2), the attributes are foot length, sex (male). In Figure 5, it is the tree for leisure shoes. Same here for 1) and 2), the foot length is the major attribute for decision.

We develop the web application for suggesting the shoes as well. Sample user interface is shown in Figure 6 and Figure 7. In Figure 6, the user enters his profile about name, date of birth, sex. Then, in Figure 7, the user enters their feet information for all about foot attributes that were mentioned. Then, the suggested shoes are displayed for each type along with the accuracy in Figure 8.

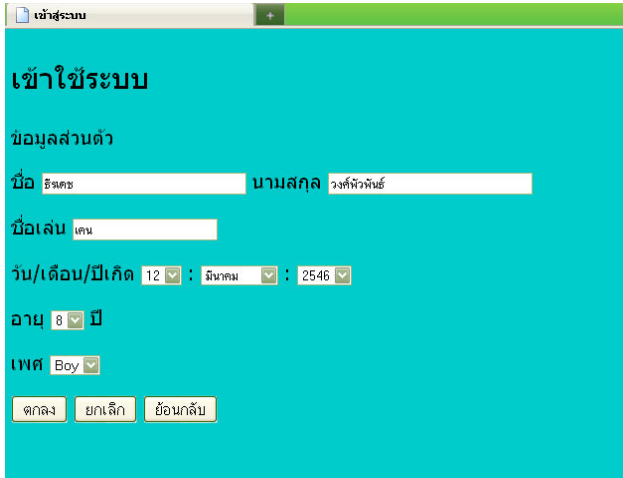


Fig. 6. User profile

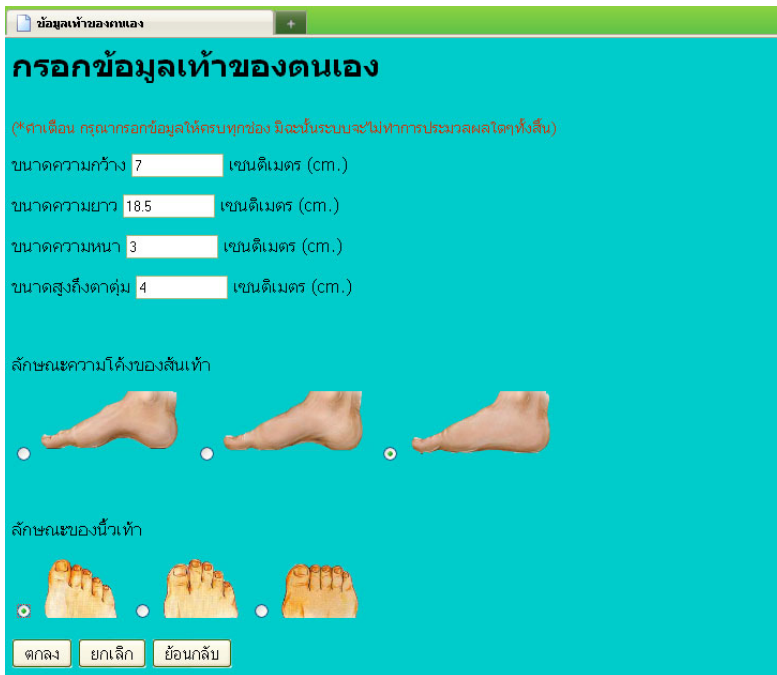


Fig. 7. User foot information

รองเท้าที่เหมาะสม คือ

รองเท้าหนังเรียบ			รองเท้าผ้า			รองเท้าผ้ารอง		
ชื่อ	เบอร์	% ความเหมาะสม	ชื่อ	เบอร์	% ความเหมาะสม	ชื่อ	เบอร์	% ความเหมาะสม
 <p>รองเท้าหนังเรียบชาย ODA รุ่น Ben Ten สีดำ</p>			 <p>รองเท้ากีฬา BREAKER รุ่น 4X4 สีขาว</p>			 <p>รองเท้าแตะ BATMAN (ลิขสิทธิ์)</p>		
	26	100		35	100		32	75

Fig. 8. Suggested shoes with accuracy

5 Conclusion

In this paper, we use a data mining to help suggest shoes for children. We are interested in factors that make the children select appropriate shoes for their comfort. The studied attributes include foot length, width, sole curve, foot style, as well as user profiles. However, the collected data may not be sufficient to train the model to see the affect of the shoe selections. Currently, the length of the shoes is the major concern. Due to this, decision tree performs well among the three approaches. In reality, sole curve, width, thickness are important to the comfort as well. To reflect this, data about the shoes need to be collected more intensively and in more detailed. This will be investigated further.

We also develop a web application that integrates the model. The user can enter their foot information and the system will suggest the proper shoes with accuracy.

References

- [1] Aitkenhead, M.J.: A co-evolving decision tree classification method. Journal Expert Systems with Applications: An International Journal 34(1) (January 2008), doi:10.1016/j.eswa.2006.08.008
- [2] Chen, Y.L., Chen, J.M., Tung, C.W.: A data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales. Journal Decision Support Systems 42(3) (December 2006), doi:10.1016/j.dss.2005.12.004
- [3] Hsieh, C., Huang, S.: Mining product maps for new product development Expert Systems with Applications 34(1), 50–62 (2008)

- [4] Hudson, S., Ritchie, B.: Understanding the domestic market using cluster analysis. Journal of Vacation Marketing 8(3), 263–276 (2002), doi:10.1177/135676670200800305
- [5] Jiao, J.R., Zhang, Y., Helander M.G.: A Kansei mining system for affective design. Expert Syst. Appl., 658–673 (2006)
- [6] Ma, C.Y., Buontempo, F.V., Wang, Z.: Inductive Data Mining: Automatic Generation of Decision Trees from Data for QSAR Modeling and Process Historical Data Analysis. In: 18th European Symposium on Computer Aided Process Engineering, vol. 25(Idm), pp. 581–586. Elsevier (2008)
- [7] Palaniappan, S., Awang, R.: Web-based Heart Disease Decision Support System Using Data Mining Classification Modeling Techniques. In: Proceedings of iiWAS 2007, pp. 157–167 (2007)
- [8] Prassas, G., et al.: Recommender System for Online Shopping Based on Past Customer Behaviour. In: e-Everything: e-Commerce, e-Government, e-Household, e-Democracy: The 14th Bled Electronic Commerce Conference, Bled, Slovenia, June 25–26 (2001)

Appendix

แบบฟอร์มการกรอกข้อมูลของเท้า


ชื่อ(ต.ช./ค.ญ.).....นามสกุล.....


วัน/เดือน/ปีเกิด.....อายุ.....ปี


ระดับชั้น.....โรงเรียน.....


คำชี้แจง: ขอความกรุณาผู้ปกครองกรอกข้อมูล และวาดรูปเท้าของนักเรียน (แบบไม่ใส่ถุงเท้า) โดยตามท่าสองบนกระดาษและใช้ดินสอในการวาด
 รองเท้านักเรียน เบอร์..... นิ้ว..... รองเท้ากีฬา เบอร์..... นิ้ว..... รองเท้าแตะ เบอร์..... นิ้ว.....


ทำเครื่องหมาย ใน
 ที่ตรงกับลักษณะของตนเอง




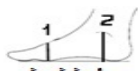











 1 2
 ว่างเท้านักเรียน

ความสูง 1 = ซม.

ความสูง 2 = ซม.

เสร็จแล้ว

เสร็จ ยาววัดด้านหลัง →

Proper Parallel Numerical Integration Method for Magnetic Sounding Equations on Multilayered Earth

Banpot Dolwithayakul¹, Chantana Chantrapornchai¹, and Suabsakul Yooyeunyong²

¹ Department of Computing, Faculty of Science, Silpakorn University, Meaung,
Nakhonpathom, Thailand

banpot@su.ac.th, chantrapornchai@gmail.com

² Department of Mathematics, Faculty of Science, Silpakorn University, Meaung,
Nakhonpathom, Thailand

suabkul@su.ac.th

Abstract. In developing a visualization of earth modeling, significant computation lies on integral equations since they come from very high oscillation. We develop the parallel algorithm for this integral equation solver based on Message Passing Interface (MPI) and OpenMP. The solver is based on the job scheduler model to balance loads for each node with threads. In the paper, we also investigate the suitable integration method for magnetic sounding equations. We have chosen five approximation methods for numerical integrations with fixed relative errors. We found that Gaussian Quadrature is the best approximation method for parallel integration of magnetic sounding equation.

Keywords: Cluster computing, High performance computing, Numerical integration, Magnetic sounding.

1 Introduction

In this research, we are interested in modeling cracks in Earth. Lots of integral equations need to be solved. These integral equations are special since they come from high oscillation, particularly magnetic sounding numerical integration.

In this paper we are attempt seeking for proper integration method for our magnetic sounding equations. There's several favorite methods for integration which we may find it useful. From our recent researches we only pick the one method "Gaussian quadrature" for integral approximation. We have not yet compared with other numerical integration methods. This paper should make us decide which method is really suitable for parallel integration of magnetic sounding equations.

In 1991, Edwards and Nabighian has built a mathematical model for n-layers earth[5]. Each layer corresponds to a unique constant value of electromagnetic conductivity except the last layer which has an infinite depth. The research is used to find the average electromagnetic field. The work has been used for finding natural resources underground successfully.

In 2000, Siew and Yooyeunyong had rebuilt the mathematical model by using electromagnetic conductivity[7]. They measured the electromagnetic field based on the frequency domain with Syscal Kid measuring device. This work can identify the mineral in a disc shape which is underground on the 2-layered earth and the mineral can be found on the lower layer. On the upper layer, the magnetic conductivity is in form of exponential equations while the lower layer has very low magnetic conductivity. To solve these exponential equations, high performance computing is needed.

1.1 Magnetic Sounding Equations

For a mathematical model, we use Maxwell’s equations[3] in Equations (1) and (2)

$$\nabla \times \vec{E} = \vec{0} \tag{1}$$

$$\nabla \times \vec{H} = \sigma \vec{E} \tag{2}$$

From Eq. (1) and (2) we can derive Equation (3).

$$\nabla \times \frac{1}{\sigma} \nabla \times \vec{H} = \vec{0} \tag{3}$$

Here H is the magnetic field density vector and Sigma(σ) is an arbitrary constant which specifies conductivity of medium. On earth modeling we use σ to specify the conductivity of the medium in the earth layers which tells us know the kind of the material occupied on each earth layer. The basic mathematical models for magnetic sounding are partial differential equations. It is written in a cylinder coordinate as in Equation (4).

$$\frac{\partial^2 H}{\partial z^2} + \sigma \frac{\partial}{\partial z} \left(\frac{1}{\sigma} \right) \frac{\partial H}{\partial z} - \frac{\partial^2 H}{\partial r^2} + \left(\frac{1}{r} \right) \frac{\partial H}{\partial r} - \left(\frac{1}{r^2} \right) H = 0 \tag{4}$$

In this equation, r is the distant between electromagnetic field generator and measuring device. This differential equation is difficult to solve with few known values. Then, we use Hankel transformation to simplify this equation.

Hankel transform of order ν of a function $f_\nu(k)$ is given by Equation (5).

$$F_\nu(k) = \int_0^\infty f(r) J_\nu(kr) r dr \tag{5}$$

We transform Equation (4) using Equation (5) and we use \tilde{H} for Hankel transformed function. The transform function shown in Equation (6) and can be transformed back as Equation (7).

$$\tilde{H}(\lambda, z) = \int_0^\infty \lambda r H(r, z) J_1(\lambda r) dr \tag{6}$$

$$H(r, z) = \int_0^\infty \tilde{H}(\lambda, z) J_1(\lambda r) d\lambda \tag{7}$$

In this equation, $J_n(x)$ is the bessel function of order n which can be estimated by the integral function in Equation (8).

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(n\tau - x \sin \tau) d\tau. \tag{8}$$

The solution for Equation (6) for k^{th} -layer where $1 \leq k \leq n$ and $n \geq 2$ is [4,7,8].

$$\tilde{H}_k(r,z) = A_k e^{-\lambda z} + B_k e^{\lambda z} \tag{9}$$

We put Equation (7) back to Equation (5) to get the final Equation (10)[7].

$$H_k(r,z) = \int_0^\infty (A_k e^{-\lambda z} + B_k e^{\lambda z}) J_1(\lambda r) d\lambda \tag{10}$$

In the equation, A and B are the arbitrary constants. This function is in the form of very high oscillation. We implement the parallel integration based on the above 5 numerical integration methods using MPI and OpenMP[6,11].

2 Integration Methods

We use the main 5 methods of integration with the fixed relative error.

2.1 Rectangle Method [16]

The rectangle method is one of the approximation of finite integral by computing area of each rectangle where its high is the value of function at that point. This method can be written as Equation (11),

$$\int_a^b f(x) dx \approx \sum_{i=1}^n f(a + i' \Delta) \Delta \tag{11}$$

where a and b are starting point and end point respectively. Δ is for $(b - a)$ and i is the variable for approximate the top corner of the rectangles. Its error(E) can be defined as Equation (12).

$$E \leq \frac{(b - a)\Delta^2}{24} f''(\xi) \tag{12}$$

For this research we use midpoint approximation for this method

2.2 Composite Trapezoidal Rule [17]

The trapezoidal rule is used for approximating the area under the function value graph as trapezoid. It can be written as Equation (13).

$$\int_a^b f(x) dx \approx \frac{b - a}{N} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{N-1} f\left(a + k \frac{b - a}{N}\right) \right]. \tag{13}$$

Its error can be define as

$$\text{error} = -\frac{(b-a)^3}{12n^2} f''(\xi), \tag{14}$$

2.3 Simpson’s Rule

Simpson’s Rule is a numerical integration method by using the following equation for an approximation as in Equation (15).

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \tag{15}$$

Its error can be approximate as (16)

$$\frac{(b-a)^5}{2880} \left| f^{(4)}(\xi) \right|, \tag{16}$$

2.4 Simpson’s 3/8 Rule [18]

There is another method for Simpson by using the cubic interpolation. This can be written as Equation (17).

$$\int_a^b f(x) dx \approx \frac{3h}{8} \left[f(a) + 3f\left(\frac{2a+b}{3}\right) + 3f\left(\frac{a+2b}{3}\right) + f(b) \right] = \frac{(b-a)}{8} \left[f(a) + 3f\left(\frac{2a+b}{3}\right) + 3f\left(\frac{a+2b}{3}\right) + f(b) \right] \tag{17}$$

The error for this method is displayed in Equation (18).

$$\left| \frac{(b-a)^5}{6480} f^{(4)}(\xi) \right|, \tag{18}$$

2.5 Gaussian Quadrature [7,8]

Gaussian is used for the approximation of integral by using “weighted sum” of function values in the domain of integration. This can be written as Equation (19).

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^n w_i f(x_i). \tag{19}$$

The weight function can be computed by Equation (20).

$$W(x) = (1-x^2)^{-1/2} \tag{20}$$

Its error can be estimated by

$$\frac{(b-a)^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\xi). \tag{21}$$

In our experiments, we will use these five methods for parallel numerical integration of Equation (10) with varied relative error on 32-node 64 cores cluster using MPI and OpenMP.

3 Parallel Integration Strategy

Since the computation we involved is the integral finding part, we deal with one dimensional array data. The intuitive way to divide the job is by given a fixed chunk of array for each node to process. Due to the subtlety of the equations, we need to find a proper integration range. This affects the job size. We discuss the issues as following.

Since Equation(9) is in the infinite range, we need to find the end of the integration range. The magnetic equation is always converged to zero.

We split the integral range into chunks. We use a rectangle method for quick and rough integration to find the end of integral range. We integrate 20 values each round until we get the absolute value of the result less than our predefined value We picked 20 because it is efficient for our sample data which we show in our recent research[2].

We use the Bessel estimation function from GNU Scientific Library (GSL) [12] to estimate Bessel function in Equation (10). The node with rank 0 separates the whole job and sends a chunk range data using blocking communication by MPI_Send to other nodes. Then, each node starts its computation. When each node finishes the computation, the root node uses a collective operation MPI_ALLREDUCE with operator MPI_SUM to combine results from other nodes. The maximum speedup for fixed-size job distribution is 18.03 on a 32-node cluster.

At the first time we separated job size to each node evenly. But we found that workload is not balance. We use the job distributor approach for solving this issue.

After we get a finite integration range, we select node 0 as a job distributor node. We use this node to divide all jobs evenly as a smaller chunk and create a data package that is sent to other nodes.

In particular, the distributor node first starts sending a start integral range to each node sequentially. The receiver node will do computation for its range, after the computation is finished, it will send the result back to the distributor node and requests a new integral range from it if one is available.

When no more integral range for distribution on distribution node, the computation will be finished after all nodes send results back to distributor node.

However, since the job distribution is a lightweight process, we create 2 threads on the distribution node. The first one will be use for normal calculation as same as other nodes and another thread will be used for the job distribution. We use OpenMP library for the thread management in this application.

We use multi-core CPU advantages by using threads on each node. For each chunk of a job, a node receives an integral range from the job distributor. We evenly divide the integral range with the total number of threads on that node and send each divided integral range to compute on each thread.

After the computation is finished, all the threads perform the reduction operation using the sum operator. Then it sends the integration result back to the job distributor. Using this approach, we can increase overall speedup to 61.14 compared with serial computation with OpenMP.

4 Experimental Result

We perform our experiments on a 32 node/64 core Linux cluster, with a Gigabit Ethernet interconnection at Louisiana Technology University, USA. In the cluster, each core is Intel Xeon 2.8GHz with 512 MB RAM. The cluster runs LAM-MPI version 7.1

We test the integration with a fixed relative error = 10^{-10} . The average computational time (t_n) is measured in seconds for N nodes.

We define relative error(ϵ) as a fraction of the actual value, if S_{exact} and S_{approx} are the analytic and approximated integration results respectively, we can write relative error value in the equation (21)

$$\epsilon_{relative} = | (S_{exact} - S_{approx}) / S_{exact} | \tag{21}$$

For easier to understand about the relative error we show the procedure for relative error control for integration in Fig.4

In each node, we used 4 threads with OpenMP for integration because we have shown that 4 threads is the best number in our recent work. The timing result for each integration technique will show on the Table 1

Table 1. Time used for parallel integration for difference number of node(N) and relative error of 10^{-10} in seconds

N	Rectangle	Trapezoidal	Simpson’s Rule	3/8 Simpson Rules	Gaussian Quadrature
2	45211	44537	18735	19305	10515
4	21324	23410	8420	8885	6123
8	8677	9847	4221	4352	2620
16	3364	3374	1768	1687	1012
32	1097	883	674	552	340

The usage time for each method can be plotted in Fig.1

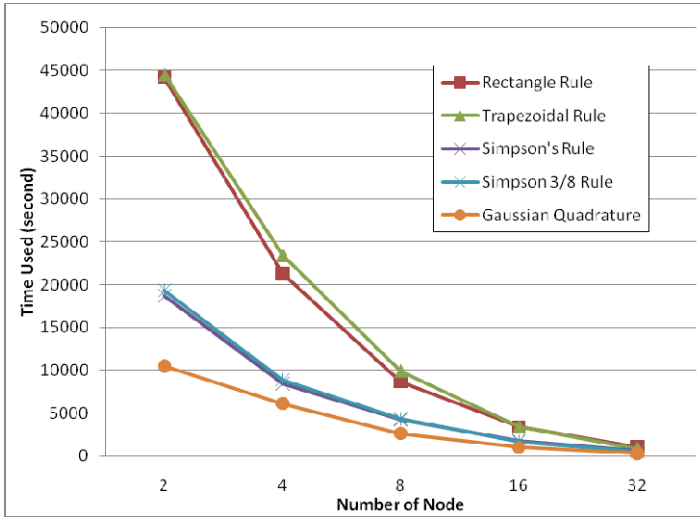


Fig. 1. Time comparison in second for various kinds of methods

As we see from the result in Table 1 we can see the Gaussian Quadrature Rule makes the best time followed by 3/8 Simpson’s Rule and Simpson’s Rule. We noticed that Gaussian Quadrature used fewer rounds for integration and has significantly lower final relative error value for integration as we show in Table 2.

Table 2. Final relative error for each method for Gaussian Quadrature 32 node integration

Method	Relative Error (ϵ)
Rectangle Rule	5.75×10^{-11}
Trapezoidal Rule	2.28×10^{-11}
Simpson’s Method	1.30×10^{-11}
3/8 Simpson’s Method	1.12×10^{-11}
Gaussian Quadrature	7.10×10^{-12}

5 Conclusion

We set the experiment for numerical integration for magnetic sounding on multilayered earth equations. Five favorite integration methods: rectangle rule, trapezoidal rule, Simpson’s Rule, 3/8 Simpson’s Rule and Gaussian Quadrature has been used for integration with fixed relative error to, 10^{-10} . The integration results were checked for correctness. We’ve found that Gaussian quadrature integration method has the best speedup for precise integration.

References

1. Dolwithayakul, B., Chantrapornchai, C., Yooyeunyong, S.: Magnetic Sounding from Multilayered Earth Equations Solver Using Job Distributor Based on Message Passing Interface. In: International Conference on Electrical Engineering/Electronics Computer and Information Technology (ECTI-CON), May 19-21, pp. 308–312 (2010)
2. Dolwithayakul, B., Chantrapornchai, C., Yooyeunyong, S.: Solving Magnetic Sounding Integral Equations from Multilayer Earth Using Message Passing Interface. In: Proceeding of International Annual Symposium on Computational Science and Engineering, pp. 454–459 (2010)
3. Maxwell, J.C.: A Dynamical Theory of the Electromagnetic Field. *Phil. Trans. Roy. Soc. London* 155, 459–512 (1865)
4. Banerjee, B., Sengupta, B.J., Pal, B.P.: Apparent resistivity of a multilayered earth having exponential varying conductivity. *Geophysical Prospecting* 28, 435–452 (1980)
5. Edwards, R.N., Nabighian, M.N.: The magnetometric resistivity method. In: Nabighian, M.N. (ed.) *Electromagnetic Methods in Applied Geophysics*, vol. 2, Part A, pp. 47–104. Society of Exploration Geophysicists (1991)
6. Bebendorf, M., Kriemann, R.: Fast parallel solution of boundary integral equations and related problems. *Comput. Vis. Sci.* 8, 121–135 (2005)
7. Yooyuanyong, S., Chumchob: Mathematical modeling of electromagnetic sounding for a conductive 3-D circular cylinder body embedded in half-space. In: Proceeding of the Third Asian Mathematical Conference, pp. 590–603 (2000)
8. Chave, A.D.: Numerical integration of related Hankel transforms by quadrature and continued fraction expansion. *Geophysics* 48, 1671–1686 (1983)
9. Jeffrey, A.: *A Handbook of Mathematical Formulas and Integrals*. Academic Press, London (1995)
10. William, H., Brian, P., et al.: *Numerical Recipes in C*, 2nd edn. Cambridge University Press (1988)
11. Quinn, M.J.: *Parallel Programming in C with MPI and OpenMP*, pp. 318–335. McGraw-Hill (2004)
12. GNU Organization, Full documentation for GNU Scientific Library (2009), <http://www.gnu.org>
13. Kim, H.S., Lee, K.: Response of a multilayered earth with a layer having exponential vary resistivity. *Geophysics* 61, 180–191 (1996)
14. Raghuwanshi, S.S., Singh, B.: Resistivity sounding on a horizontally stratified varying resistivities. *Geophysical Prospecting* 34, 409–423 (1986)
15. Rjasanow, S., Steinbach, O.: *The Fast Solution of Boundary Integral Equations*, pp. 135–160. Springer, Heidelberg (2007)
16. Shilov, G.E., Gurevich, B.L.: *Integral, Measure, and Derivative: A Unified Approach*. Dover Publications (1978); Silverman, R.A. (trans.)
17. Weideman, J.A.C.: Numerical Integration of Periodic Functions: A Few Examples. *The American Mathematical Monthly* 109(1), 21–36 (2002)
18. Matthews, J.H.: Simpson's 3/8 Rule for Numerical Integration, Numerical Analysis - Numerical Methods Project. California State University, Fullerton (2004), <http://math.fullerton.edu/mathews/n2003/Simpson38RuleMod.html>

QoS Guarantee with Adaptive Transmit Power and Message Rate Control for DSRC Vehicle Network Based Road Safety Applications

Wenyang Guan¹, Jianhua He², Lin Bai^{3,*}, Zuoyin Tang², and Yi Zhou¹

¹ College of Engineering, Swansea University, UK

² School of Engineering and Applied Science, Aston University, UK

³ School of Electronic and Information Engineering, Beihang University, China
j.he7@aston.ac.uk, l.Bai@buaa.edu.cn

Abstract. Quality of services (QoS) support is critical for dedicated short range communications (DSRC) vehicle networks based collaborative road safety applications. In this paper we propose an adaptive power and message rate control method for DSRC vehicle networks at road intersections. The design objective is to provide high availability and low latency channels for high priority emergency safety applications while maximizing channel utilization for low priority routine safety applications. In this method an offline simulation based approach is used to find out the best possible configurations of transmit power and message rate for given numbers of vehicles in the network. The identified best configurations are then used online by roadside access points (AP) according to estimated number of vehicles. Simulation results show that this adaptive method significantly outperforms a fixed control method.

1 Introduction

Road traffic safety has been a subject of worldwide concern. Driven by advances in wireless communications and mobile networking, collaborative safety applications (CSA) enabled by vehicular communications are widely considered to be important to future road safety [1]. DSRC is widely regarded as the most promising technology to provide short range vehicular communications. The technology is expected to be robust and affordable enough to be built into large-scale vehicles [2][3][4]. DSRC standards are currently being developed by organizations including the IEEE [5]. IEEE is specifying a wireless access in vehicular environment (WAVE) for DSRC to provide seamless, interoperable V2V and vehicle to RSU (V2R) communication services for transportation [5].

Despite the great potentials of DSRC technology for CSA, there are still many challenges faced by the practical deployment of DSRC based CSA systems. One challenge is that road safety applications have stringent requirements for reliable real-time message delivery, as excessive message delays or message loss hinder the effectiveness of CSA and can even cause unexpected negative consequences.

* Corresponding author.

However, as the MAC layer of DSRC is based on the IEEE 802.11 distributed coordination function (DCF) [6], the stringent service requirements is hard to be met by the random channel access specified in the IEEE 802.11 DCF [7]. Provisioning of QoS for the road safety applications especially for ESA is critical for CSA [8] [9].

In this paper we investigate the issue of QoS support from DSRC vehicle networks based on message rate and power control. Two types of collaborative road safety applications are considered for QoS support, which are assumed to be deployed over the DSRC control channel (CCH) [10]. One is event-driven safety applications (ESA), and the other is periodic safety applications (PSA). ESA is designed to be used for emergency scenarios. It generates emergency messages if accidents happen or are emerging. ESA messages have high priority and timely and robust communication services. Position information obtained by GPS is normally included in the ESA messages. PSA is designed for announcing existence of a vehicle by non-emergent messages. PSA messages are periodically generated and broadcasted to build mutual awareness. For example the basic information can be carried by PSA messages, such as velocity, acceleration or deceleration [1].

The main idea of the proposed adaptive control method is that we use an off-line simulation based approach to find out the best possible message rate and transmit power for given numbers of vehicles. The identified best configurations are then adaptively used online by roadside APs to control the network congestion and QoS. The authors have studied adaptive message rate control for DSRC vehicle networks, in which safety message rate is controlled in a distributed manner by the vehicles in a freeway [9]. Centralized message rate control for intersection road is proposed but transmit power control is not considered. It is noted without transmit power control the adaptive control scheme proposed in [11] could provide a certain degree of QoS but could not guarantee the required QoS for safety applications. To the best of our knowledge, our work presented in this paper is first of its kind on QoS guarantee for DSRC vehicle networks based collaborative road safety applications.

2 Adaptive Congestion Control

As we mentioned earlier the primary design objective is to provide high availability and low latency channel for high priority ESA messages. The secondary objective of this design is to maximize channel utilization for low priority PSA messages. There are two major parts included in the proposed adaptive control method. The first part is an offline procedure to find out the optimal configurations of message rate and transmit power for a set of given number of vehicles. These optimal configurations are then applied in the second part where the roadside AP requests the vehicles to update the configuration of message rate and BE according to an estimated number of vehicles in the road intersection.

2.1 Offline Determination of Optimal Configurations

The offline procedure is to find out the optimal configurations of message rate and transmit power. We implement a simulator to determine the optimal configurations for each given number of vehicles. A challenge on the determination of optimal configuration is the multiple objectives optimization for the whole vehicle network. For example, a low PSA message rate will leave higher channel availability to ESA messages at the cost of possibly low number of PSA messages transmitted successfully.

To tackle the multiple objective optimization problem we use a utility function to measure the overall network performances. In the proposed utility function the performance metrics of message success probability and average transmit delay are taken into account. Let $P_{s,e}$ and $P_{s,p}$ denote message success probability for ESA messages and PSA messages, respectively. Let D_e and D_p denote message delivery delay for ESA messages and PSA messages, respectively. Let R_e and R_p denote selected message rate for ESA messages and PSA messages, respectively. The proposed utility function (denoted by Θ) can be expressed by:

$$\begin{aligned} \Theta = & R_e W_{s,e} (P_{s,e} - P_{thr,e})^+ + R_p W_{s,p} (P_{s,p} - P_{thr,p})^+ \\ & - W_{d,e} (D_e - D_{thr,e})^+ - W_{d,p} (D_p - D_{thr,p})^+ \end{aligned} \quad (1)$$

where $P_{thr,e}$ and $P_{thr,p}$ are preset thresholds for message success probability of ESA and PSA messages, respectively. $D_{thr,e}$ and $D_{thr,p}$ are preset thresholds for delivery delay of ESA and PSA messages, respectively. $W_{s,e}$ and $W_{s,p}$ are preset utility weights for message success probability of ESA and PSA messages, respectively. $W_{d,e}$ and $W_{d,p}$ are preset utility weights for delivery delay of ESA and PSA messages, respectively. The threshold function $(x)^+$ used in the utility function is expressed by

$$(x)^+ = 0, \text{ if } x \leq 0 \quad (2)$$

$$(x)^+ = x, \text{ if } x > 0 \quad (3)$$

With the utility function and the preset parameters we can set an optimal configuration table for message rate and BE which maximizes the utility function Θ for a given number of vehicles in the vehicle network. Note that the determination of the optimal configurations is to be done by the AP only.

2.2 Online Adaptive Configurations

In this procedure the AP applies the findings from the offline procedure on the optimal configurations according to estimated number of vehicles in the road intersection. Firstly the AP estimates the number of vehicles (N_{est}) in the road intersection 60 seconds through the received PSA messages. According to the estimated number of vehicles the AP looks up the optimal configuration table to get the optimal configuration for the preset QoS requirements.

The optimal configuration is then broadcasted in ESA messages by the AP to the vehicles in the intersection. Vehicles received the AP ESA messages update their configurations according to the instructions received from the AP.

3 Numerical Result

A discrete event driven simulator has been developed to evaluate the adaptive congestion control method for DSRC vehicle networks. We assume a single hop ad hoc network environment without hidden terminals. For the DSRC control channel we set CCH interval as 35 ms, which is 70% of a 50 ms Synchronization Interval (SI). For simplicity we assume there are two classes of vehicle nodes in the network. The first class of nodes only transmit ESA messages while the second class of nodes transmit only PSA messages. In the simulations we simply assume that there are three first class nodes, which send eight ESA messages in each second. All messages are assumed to have the same length (250 bytes) and are broadcasted at the rate of 3 Mbps. An ideal channel is assumed, which means a message can be successfully received if no collision happens.

In this paper we have used the following configuration for the thresholds and utility weights for the proposed utility function. $P_{thr,e} = 0.9$, $P_{thr,p} = 0.8$, $W_{s,e} = 0.7$, $W_{s,p} = 0.3$, $W_{d,e} = 10$, $W_{d,p} = 20$. With the above parameter configurations we set up the optimal configuration table for message rate and transmit power. The utility Θ for the optimal configurations is plotted in Fig. 1 against the number of vehicle nodes in the network. For performance comparison purpose the system utility Θ obtained with a fixed control scheme which uses fixed message rates and transmit power is also presented in Fig. 1. The corresponding message rate R for PSA application which maximize the utility function are presented in Fig. 2. We plot in Fig. 3 the message delivery delay for both ESA and PSA messages obtained with the optimal configurations.

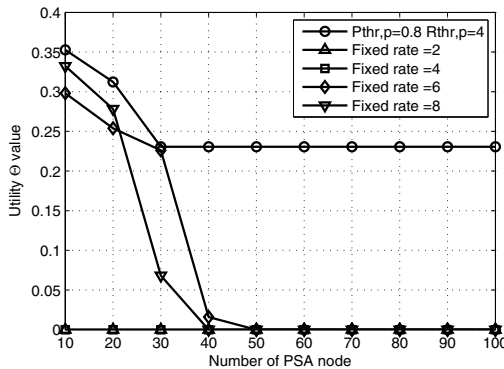


Fig. 1. System utility Θ with optimal configuration and fixed configuration against the number of PSA nodes

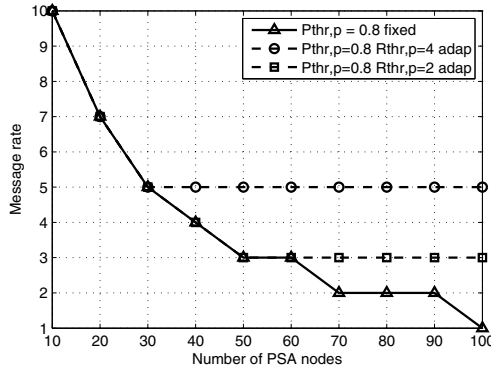


Fig. 2. Optimal message rate against the number of PSA nodes, for $P_{thr,p} = 0.8$

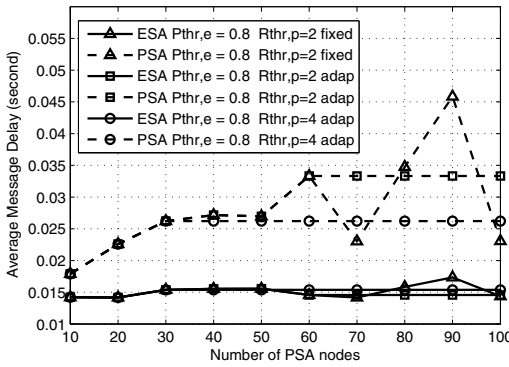


Fig. 3. Average message delivery delay

It can be observed that the performance with optimal configuration is consistently significantly better than that with fixed rate and transmit power. And more specifically when the number of vehicles in the network is high, controlling message rate alone could not guarantee the required QoS for safety applications. In these cases the proposed adaptive control scheme could provide guaranteed QoS with joint control of message rate and transmit power.

Acknowledgement. The work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) with grant reference number EP/1010157/1, the China National “973” project under the grant No.2009CB320403 and the National Natural Science Foundation of China (NSFC) under the grant number 61103177.

References

1. Vehicle Safety Communications Project - FINAL REPORT. CAMP IVI Light Vehicle Enabling Research Program, DOT HS 810 591 (2006)
2. Zhu, J., Roy, S.: MAC for dedicated short range communications in intelligent transportation system. *IEEE Commun. Mag.* 41, 60–67 (2003)
3. Biswas, S., Tatchikou, R., Dion, F.: Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety. *IEEE Commun. Mag.* 44, 74–82 (2006)
4. Jiang, D., et al.: Design of 5.9 GHz DSRC-based vehicular safety communication. *IEEE Wireless Commun.*, 36–43 (October 2006)
5. IEEE 802.11p, “Amendment for Wireless Access for the Vehicular Environment (WAVE)” (2010)
6. IEEE Std. 802.11e, Wireless LAN Medium Access Control (MAC) Enhancements for Quality of Service (QoS) (2005)
7. Yin, J., Elbatt, T., Habermas: Performance evaluation of safety applications over DSRC vehicular ad hoc networks. In: *ACM VANET 2004* (2004)
8. He, J., Tang, Z., O’Farrell, T., Chen, T.: Performance analysis of DSRC priority mechanism for road safety applications in vehicular networks. In: *Wireless Commun. and Mobile Comp.* (August 2009)
9. He, J., Chen, H.-H., Chen, T., Cheng, W.: Adaptive congestion control for DSRC vehicle networks. *IEEE Commun. Letters* 14, 127–129 (2010)
10. IEEE Std 1609.4, “IEEE Trial-Use Standard for Wireless Access in Vehicular Environments (WAVE) Multi-channel Operation” (2006)
11. Guan, W., He, J., Bai, L., Tang, Z.: Adaptive Congestion Control of DSRC Vehicle Networks for Collaborative Road Safety Applications. In: *IEEE Workshop on Wireless Local Area Networks, WLN 2011* (2011)

The Influence of Wine Expert's On-Line Ratings on Retail Price of On-Premise Market in Korea

YoonJung Nam¹ and Youngsik Kwak²

¹ Ph. D. Candidate, Dept. of Tourism, Hanyang University
jean@hanyang.ac.kr

² Associate Professor, Gyeongnam National University of Science and Technology,
#150, Chilam-dong, Jinju, Korea
yskwak@gntech.ac.kr

Abstract. Quality of wines can be evaluated in 4 different ways; 1) Subjective evaluation by experts, 2) Subjective evaluation by consumers, 3) Objective evaluation by rating agencies on a regular basis, and 4) Objective evaluation by rating agencies on a temporary basis. The purpose of this study is to explore the influence of subjective evaluation on quality of wines by wine experts on retail prices in Korean wine market. This study focused on empirically the influence of 'Robert Parker Scores' which are revealed on a website owned by Robert Parker Jr., known as the most influential wine critic in the world on retail prices of 1,039 wines produced in 3 different countries, and sold in 7 restaurants of hotels and wine bars in Korea. Furthermore, this study empirically analyzed relative importance of influences that vintage data and Parker Score affect on retail price of on-premise market.

As a result, this study extended previous research on influence of experts' evaluation result (Parker Score) on wholesale price level to retail price level. Furthermore, this study also extended previous researches on influence of Parker Score on Bordeaux wines only to wines produced in various wine producing regions of France, Italy, and U.S.A: Champagne, Burgandy, Rhone, Piemonte, Toscana, and California..

Keywords: Wine Industry, on-line critics, Robert Parker, Retail Price.

1 Introduction

Many studies on relationship between price and quality have been done in marketing research and these studies can be categorized into 2 types: 1) relationship between price and subjective quality rated by consumers, 2) relationship between price and objective quality [1].

Among these, most of studies on price-subjective quality have been done with various categories of products and services in cross-sectional studies using price as an indicator of quality. This kind of studies started in early '70s and has been made for 40 years[2] [3] [4] [5] [6] [7]. In respect of evaluating, there are two performers; individual consumers and experts. Most of studies have been done on relationship between price and subjective quality rated by individual consumers, whereas few studies on price-subjective quality rated by experts have been done.

On the other hand, third-party expert evaluate products quality in the studies on correlation of price-objective quality. These kinds of studies have been performed with various subjects in various countries including Japan, Korea, US, and Canada.[8] [9] 10] At these studies, evaluation is performed by organizations or agencies provide evaluation results on regular bases or on temporary bases.

Therefore, studies on price-quality correlation can be categorized into 4 types. The first one is studies discussing on price-consumer's subjective quality, in other words, on wine price-wine consumer's subjective quality ratings. However, this type of studies are not easily found in the field of wine industry.

The second one is related to the price-expert's subjective quality. Studies discussing on influence of Parker Score affect on the price of en-primeur can be categorized into this type of studies [11][7]. However, these studies have limitation with focusing on Bordeaux wines of certain vintages of 1994-1998, 2003, 2005, and 2008.

The third one is belonged to studies discussing on price-objective quality rated by agencies in regular bases. In other product categories, this type of studies has been made in various countries such as USA, Europe, Japan, and Korea. For example, Consumer Report published in USA has evaluated cars to revealed objective quality ratings. In Korean market, Consumer Times plays the same role with Consumer Report's. Wine Scores and Top 100 wines of the year revealed regularly by Wine Spectator can be included in this case.

The fourth one is related to the studies discussing on the relationship between price and objective quality rated by 'temporary work agencies' such as studies on the relationship between price or Box Office and award-winning artworks or films [12]. Studies on the influence of award-winnings such as medal-awarded in international competition can be an example of this type. As empirical studies on the relationship-between price and quality ratings are not accumulated academically, this type of studies could be meaningful in wine studies and for wine industry as well.

Table 1. Previous studies on price-quality rating and position of this study

category of quality rating	main units for ratings	types of ratings	in wine industry	position of this study
subjective quality	experts	critic, on-line report, publicity(Michelin guides)	Parker Score	X
	individual consumers	posting reply, answering to survey, WOM		
objective quality	rating agency (regular basis)	Consumer Report, Consumer Times	Wine spectator's score	
	rating agency (temporary basis)	Biennale, International film festival, Competition	wine competition	

The purpose of this study is to see how 'Parker Score', wine quality ratings made by Robert Parker who is a very influential and a expert wine critic, affects on retail prices in Korean wine market. With this study, we will extend the scope of study object to various wines produced in various regions while previous studies focused on Bordeaux wines. Furthermore, while previous studies focused on the relationship between wholesale price and Parker Score, we will observe the affect made by Parker Score on retail list price in Korean wine market. This study thus aims at extending the origin of wines as a study object and also extending scope of study on pricing in wine industry from wholesale prices level to retail prices level. This study also can be a guideline advises wine origins and vintages that consumers need to consider in consuming wines to avoid overspending and also to take a chance to secure consumer surplus in Korean on-premise market.

2 Literature Review: Studies on the Relationship between Price and Subjective Quality Ratings from Experts in Wine Industry

Previous studies on the relationship between price and experts' subjective quality ratings are focused on price on en primeur - Parker Score [13][14][15]. Hay (2007) suggests that Parker has a significant impact on Bordeaux wine's en primeur price formation in appellations. Dubois and Nauges (2007) found that Parker Score had made for 5 years (1994-1998) affected on price formation of Bordeaux wines from ten different appellation in estimating significance with structural equation modeling and OLS modeling[17].

Ali et al. (2008) found and estimated a Parker effect event at unusual situation in 2003 when the wine grades were published much later, in the autumn, after the determination of prices. Hay (2010) examined the impact of the Parker Score on the process of price formation through a comparison of the 2005 and 2008 en primeur campaigns.

The year 2005 is a truly exceptional vintage, 2008 probably rather less so; 2005 has been unanimously acclaimed as an exceptional year, 2008 could scarcely be more different with hard time for harvest. This study found that the impact of the Parker Score was independent to specific conditions of the vintages.

These previous studies have some of limitations and implications for future study. First, previous studies focus on Bordeaux wines and certain vintages. Most of studies focus on vintages of 2003 and 2005 except a study of Dubois and Nauges(2007). Secondly, these studies discuss on wholesale prices in Bordeaux's en primeur market only while many studies in other product category have price variables with list prices, that is, retail price level.

Therefore, extending origins of wines evaluated subjectively by experts can be meaningful and study on impact of experts' subjective quality ratings on retail price can provide Korean wine industry with empirical implication.

3 Methodology

This study aims at filling up the blank of study on retail price in wine industry. For this, we will examine the impact of Robert Parker's subjective quality ratings on the wine retail price at regions of origin level.

Research question 1 : Does the Parker Score impact on wine retail price in Korean on-premise market?

Research question 2 : Does the Parker Score impact on retail price of wines from certain sub-regions(cf. Champagne, Bordeaux, Bourgogne, Rhone, Toscana, Piemonte, and California) of certain countries(France, Italy, and USA) in Korean on-premise market?

Research question 3 : Relative importance of the Parker Score and vintage on retail prices of wines from 7 regions of 3 countries

3.1 Measuring the Variables

Our data set combines two sources of information: retail prices of 6 hotel restaurants and 1 wine restaurant located in Seoul and Parker Score extracted from Parker's website, www.erobertparker.com These pieces of information can also be found on The Wine Advocate. 7 restaurants' wine list have 2,210 items of wines and among them, we selected 1,089 items that have Parker Score.

3.2 Sample

The Parker Score distributes from 65 to 100 and wines from 10 countries have Parker's subjective ratings in our sample. The sample has 44 wine producing regions and 32 appellations with the Parker Score. And there're 40 different vintages that have the Parker's Score. Price distributes from 17,000 KRW to 10,000,000 KRW. To make the purpose of this study, we selected 3 countries which had more than 30 samples in every sub-region (appellation). Throughout this process, 1,039 samples were finally selected.

4 Result of Empirical Analysis

4.1 The Influence of Parker Score on Retail Price without Regard to Country of Origin

Research question 1 is to find whether Parker Score affect on retail price in Korean on-premise wine market or not. For this, we set the retail price on the wine list as the dependent variable and appellation, vintage, Parker Score as the independent variable and conducted regression analysis. As a result of the analysis, determinant coefficients was 0.622 and F-value was 42.11($P < .001$).

As you can see on column of total sample on Table 2, the Parker Score and 24 vintages from 40 vintages are significant variables that impact on retail price in significant level ($P < .001$). Appellation variable was found as not significant. Parker Score shows higher relative importance in comparison 22 vintages among 24. However, vintage impact more on retail price than Parker Score does with 1982 and 1986, two of the best Bordeaux vintages of the last 30 years.

4.2 The Influence of Parker Score and Vintage Affect on Retail Price in Appellations

Research question 2 is to explore whether Parker Score and vintage affect on the price of wines produced in 7 appellations (Champagne, Bordeaux, Bourgogne, Rhone, Toscana, Piemonte, and California) of 3 countries. Research question 3 is to investigate relative importance of influence that Parker Score and vintage affecting to the retail price in Korean on-premise market. To study this issue, we conducted regression analysis with 7 wine restaurants' retail price as a dependent variable and vintage and Parker Score as independent variables. Bordeaux's adjusted R² is 0.713 which is the highest score among that of 4 appellations in France; Champagne, Bordeaux, Bourgogne, and Rhone and the appellation with lowest score is Bourgogne with 0.098. Champagne and Rhone are relatively much affected by Parker Score than Bordeaux and Bourgogne.

We assume that as there are many of 'non vintage' Champagne in Champagne section of the wine lists, the Parker Score or brand image could more affect on retail price rather than vintage. Among Bordeaux wines, 15 vintages are statistically significant. Especially, vintage of 1982 and 1986 have relatively more powerful than Parker Score. Vintage frequently affect on retail price than others as you can refer to Table 2 in case of Bordeaux.

Bourgogne's adjusted R² is 0.098 which is the lowest score among that of 4 appellations in France. Parker Score's standardized regression coefficient is 0.227($p < 0.001$). 2002 vintage is only one vintage that has statistical significance and vintage affected more than Parker Score on the retail price with its vintage. In other words, we assume that there are many factors that cannot be explained with vintage and the Parker score, and Parker Score has relatively less impacted on retail price in Bourgogne wine category. In Rhone wine category, Parker Score's standardized regression estimate is 0.585($p < 0.001$). 3 vintages had statistical significance and we found that Parker Score has powerful impact on the retail price in this category. As a result, Parker Score has different levels of impact on the retail price in each appellation. And we found that Parker Score is more powerful than vintage in general with exceptions in certain vintages. In Toscana wine category, Parker Score's standardized regression coefficient is 0.557($p < 0.001$). Among Toscana wines, 1988 and 1997 vintages have statistical significance. These results show that many factors other than vintage and Parker Score are considered in marketers' pricing. In Piemonte wine category, Parker Score's impact is very powerful on pricing with vintage's additional influence. The result reveals that Piemonte wine category is most affected by Parker Score and vintage than other 6 appellation categories. In Californian wine category, Parker Score's standardized regression estimate is 0.459($p < 0.001$) and 1997 and 2000 vintages are statistically significant. This result reveals that other variables than Parker Score and vintage were considered in pricing.

5 Conclusion

The purpose of this study is to empirically analyze the influence of subjective evaluation of an expert under on-line on retail price in Korean wine market. This study is meaningful in the fact that we extended the scope of the research from wholesale price to retail price in pricing research, and also from Bordeaux to other French appellations and other countries in measuring Parker Score's influence level. The result of this study revealed that the level of Parker Score's influence is different in wine producing county and appellation. In general, Parker Score was more powerful than vintage rating except but there was an exception with certain vintages.

Wines from Champagne, Rhone, Piemonte, and Toscana were more affected by Parker Score than wines from other appellations; Bordeaux, Bourgogne, and California. Bordeaux had more cases that vintage rating affects on retail price, especially, the influence of 1997 vintage on retail price was significant in wines from all 3 countries; France, Italy, and U.S.A. The result suggest that consumers consider Parker Score in choosing Champagne and Piemonte wines, vintage in choosing Bordeaux wines, and brand in choosing Bourgogne wines respectively for satisfaction with reasonable price.

In order to expand the results, additional research will have to be done in relation to more diverse variable's effects, such as brand, country, appellation, and so on.

References

1. Hwang, S., Beak, S.: The Comparative Study on Price and Objective Quality of Apparel Products between Korea and United States. *Journal of the Korean Society of Clothing and Textiles* 26(3-4), 527–538 (2002)
2. Gardner, M.D.: Is There a Generalized Price-quality Relationship? *Journal of Marketing Research* 8, 241–243 (1971)
3. Tellis, G.T., Wernerfelt, B.: Competitive Price and Quality under Asymmetric Information. *Marketing Science* 6(3), 240–253 (1987)
4. Heisey, L.F.: Perceived Quality and Predicted Price: Use of the Minimum Information Environment in Evaluating Apparel. *Clothing and Textiles Research Journal* 8(22-28) (1990)
5. Lockshin, L.S., Rhodus, W.T.: The Effect of Price and Oak Flavor on Perceived Wine Quality. *International Journal of Wine Marketing* 5(2-3), 13 (1993)
6. Goudie, A.J., Sumnall, H.G., Field, M., Clayton, H., Cole, J.C.: The effects of price and perceived quality on the behavioural economics of alcohol, amphetamine, cannabis, cocaine, and ecstasy purchases. *Drug and Alcohol Dependence* 89(2-3) (2007)
7. Ali, H.H., Lecocq, S., Visser, M.: The Impact of Gurus: Parker Grades and En Primeur Wine Prices. *The Economic Journal* 118(529), F158–F173 (2008)
8. Bodell, R.W., Kerton, R.R., Schuster, R.W.: Price as a Signal of Quality: Canada in the International Context. *Journal of Consumer Policy* 9, 431–444 (1986)
9. Yamada, Y., Akerman, N.: Price-Quality Correlation in The Japanese Market. *Journal of Consumer Affairs* 18, 215–265 (1984)
10. Kim, Y., Kim, E.: A Study of the Correlation between the Objective Quality and the Prices in Korean Markets. *Journal of Consumer Studies* 10(3), 85–100 (1999)

11. Hay, C.: Globalisation and the Institutional Re-embedding of Markets: The Political Economy of Price Formation in the Bordeaux En Primeur Market. *New Political Economy* 12(2) (2007)
12. Nelson, R.A., Donihue, M.R., Waldman, D.M., Calbraith, W.: What's an Oscar Worth? *Economic Inquiry* 39(1), 1–16 (2001)
13. Ali, H.H., Lecocq, S.: The pricing of experience goods: The case of en primeur wine. *American Journal of Agricultural Economics* (2006) (forthcoming)
14. Ali, H.H., Lecocq, S., Visser, M.: The impact of gurus: Parker grades and en primeur wine prices. University of Toulouse, INRA working paper 2005 (2005)
15. Hay, C.: The political economy of price and status formation in the Bordeaux en primeur market: the role of wine critics as rating agencies. *Socio-Economic Review* 8, 685–707 (2010)
16. Ginsburgh, V.: Awards, success and aesthetic quality in the arts. *Journal of Economic Perspectives* 17, 99–111 (2003)
17. Dubois, P., Nauges, C.: Identifying the Effect of Unobserved Quality and Expert Reviews in the Pricing of Experience Goods: Empirical Application on Bordeaux Wine. *American Association Of Wine Economist, Working Paper* (2007)

Shipment Timing Support System Modeling for Stored Apples in Korea

Youngsik Kwak¹, Seokil Song^{2,*}, and Yoonsik Kwak^{3,*}

1 Associate Professor, Dept. of Venture and Business, Gyeongnam National University of Science and Technology, Korea

yskwak@gntech.ac.kr

^{2,3} Dept. of Computer Eng., ChungJu National University, Korea

seokil.song@gmail.com,

yskwak@cjnu.ac.kr

Abstract. Although, in academic field, various industries has its own retail price management system from the customers' point of view, the report of application of wholesale price management in practice, especially in agricultural industry, has been relatively rare from producers' perspectives. The researcher aims to try to full this gap by developing a shipment timing support system for stored apples based on the transaction quantity and the wholesale price information announced by Seoul Agricultural & Marine Products Corporation at on-line. The Shipment Timing Support System(STSS) consists of the dynamic price response function for each product, the price expectation effect, product line effect, substitutes price effect, events effect, apples' life cycle, the demand curve for apples and cost function. The resulting fluctuation of the expected sales revenue for every product they store in warehouse will be provided by the application on smart phone. The system will helpful for farmers and wholesalers to decide the time for shipment from warehouses to wholesale markets, and to increase the opportunities to produce more money in farmers pockets.

Keywords: Stored Apples, Agriculture, Shipment Timing, Product life cycle, Support system.

1 Introduction

To increase the profit, the academic research has recommended that managers are advised to set prices for their products according to customers' perceptions of product benefits and costs. It means that the product in retail should be seen through the customers' point of view and priced accordingly[1].

However, many farmers in Korea cannot involve to set retail price directly. The most of their agricultural products are sold through a few auction-based wholesale market such as Garak and Gangseo agricultural market in Korea. The wholesale price is the core parameter for farmers and wholesalers to sell or ship their goods at the wholesale markets in Korea [2].

* Corresponding author.

Although the retail price management is well documented in academic field from the customers' point of view, the report of application of wholesale price management in practice, especially in agricultural industry, has been relatively rare from farmers' perspectives. The researcher aims to try to fill this gap by developing a shipment timing decision-making support system for stored apples based on the transaction quantity and the wholesale price information announced by Seoul Agricultural & Marine Products Corporation, every day, through on-line. The system will be helpful for farmers to decide the time for shipment from warehouses to wholesale markets, and to increase the opportunities to produce more money in farmers' pockets. The previous research required that the system faithfully reflects the farmers' shipment behavior[3].

Furthermore, to decide to ship or not to wholesale market, farmers considered so many variables such as storage cost, physical loss, wholesale price, and the expected income. Therefore, the support system introduced in this paper is based on various variables such as dynamic price response function effects, price change response for product lines, the price expectation effects, seasonal effects such as Chuseok and Seol, price response function by substitutes, demand curve for apples and the cost function.

2 Modeling

The researcher develops the STSS with two modelings. One is the dynamic price response function for apples. The other is the demand curve at a given time for apples.

2.1 Static Price Response Function and Dynamic Price Response Function

The relationship between alternative prices and the resulting sales quantity is called price response function [4]. The static price response functions between retail and wholesale level show the different curve shapes according to perspectives. At retail situation from customers' view points, the lower the price, the higher the sales quantities. However, at wholesale market from the suppliers' perspectives, the higher the wholesale price, the higher the shipment volume. Therefore, the support system should be developed under the hypothesis of the positive relationship between the wholesale shipment volume and wholesale price.

Especially, farmers will decide whether they ship the apples to market or not at a given time t , depending on the previous wholesale price level p_{t-1} . As you can see at Eq. (1), we can model the time gap between p_{t-1} and q_t . Eq. (1) can be called the static price response function.

$$q_t = f(p_{t-1}) \quad (1)$$

where q_t = shipment volume at period t and p_{t-1} = wholesale price at period $t-1$.

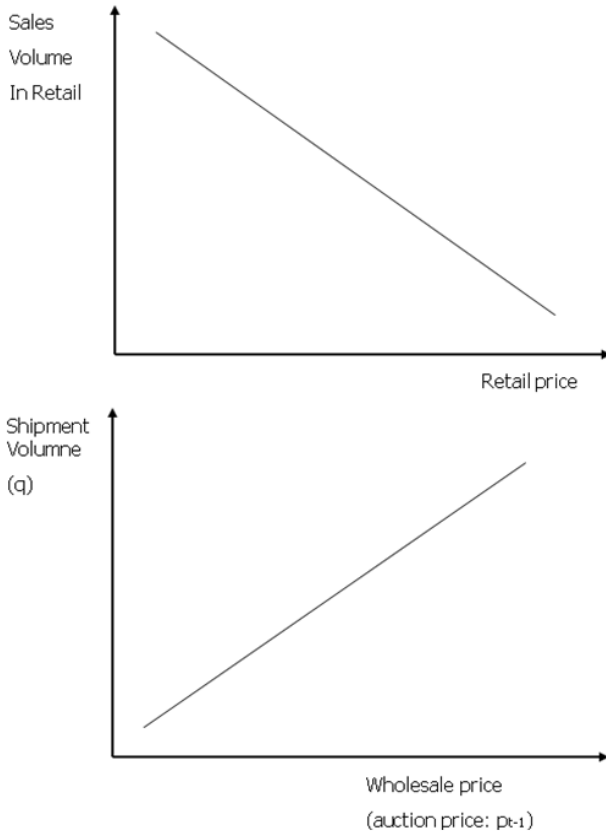


Fig. 1. The different price response function depending on channel and perspectives

Especially, farmers will decide whether they ship the apples to market or not at a given time t , depending on the previous wholesale price level p_{t-1} . As you can see at Eq. (1), we can model the time gap between p_{t-1} and q_t . Eq. (1) can be called the static price response function.

The inclusion of the time dimension requires a modification of the static price response function because the current price can have effects on future sales. Two possible effects are inferred. On the one hand, for many farmers the past wholesale prices serve as a base of price comparison and, thus, affect the current price response, that is, shipment volume. The current price assumes the same role with respect to future prices. The effect of price changes may be influenced by this initial base of comparison. Therefore, the researcher should assume the dynamic price response function at shipment volume for stored apples[5]. On the other hand, the shipment volume increase produces customers' decreasing demand. This phenomenon makes the wholesale price down, vice versa. So, the researcher includes the time dimension on the price response function. Eq. (2) can be called a dynamic price response function.

$$q_t = a t - c (p_{t-1} - p_{t-2}) \quad (2)$$

where q_t = shipment volume at period t , p_{t-1} = wholesale price at period $t-1$, p_{t-2} = wholesale price at period $t-2$ and a and c = parameters to be estimated.

2.2 Non-proportional Response to Price Changes

In many cases, it seems that small price changes cause under-proportional responses and big price changes produce over-proportional responses. For example, price reductions of less than 10% has no significant effect on sales, while the sales increase strongly when price reductions are more than 15%. This phenomenon is often observed in the sale of new products. The hypothesis of non-proportional response to price changes is similar to Gutenberg hypothesis for price differentials between competitive products[1]. Relatively, farmers who stored apple would not response to small wholes price in comparison to big price changes. Fig. 2 shows such a function with a non-proportional response to price change.

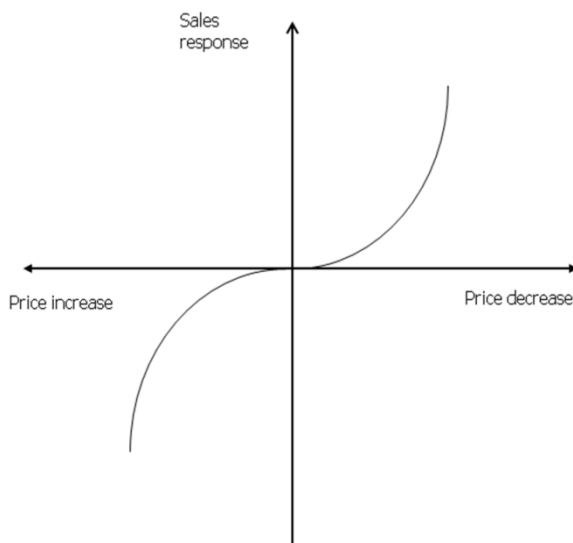


Fig. 2. Non-proportional response to price change

Formally a sinus-hyperbolic function can generate the shape of the curve. If we choose the relative price change as the independent variable, we get

$$q_t = a t - c_1 \sinh (c_2 (p_{t-1} - p_{t-2}) / p_{t-2}) \quad (3)$$

where q_t = shipment volume at period t , p_{t-1} = wholesale price at period $t-1$, p_{t-2} = wholesale price at period $t-2$ and a , c_1 and c_2 = parameters to be estimated.

2.3 Intra- and Inter-apple Pricing Effect among Complements and Substitutes

Few firms have just a single offering anymore. This multiful-product strategy presents new pricing issues and opportunities. How it affects pricing decisions hinges on the nature of interdependencies between the products [6][1]. In this regard, farmers sometimes cannot produce the same grade apple at once. Furthermore, during the storage at warehouse or cold storage facilities, the quality of apple may be down-grade according to storage circumstance. So, farmers have several apple offerings. Seoul Agricultural & Marine Products Corporation has reported 4-grade apples such as excellent, high, mediocre, and low grade, at daily base. We call apple’s four grades the apple’s product line.

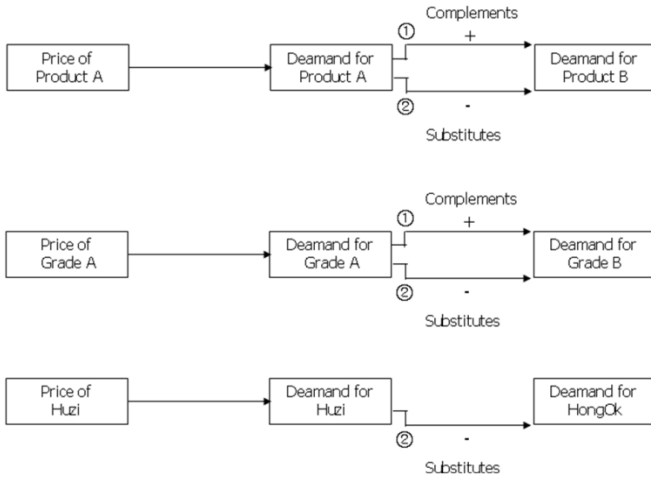


Fig. 3. Possible product-line linkages

The farmers’ product line is complements and substitutes, simultaneously. Linkages 1 and 2 in Fig. 3 show that demand for product A can directly affect demand for another product, B. If this interrelation is positive (arrow 1), i.e., if an increase in demand for A leads to higher demand for B, we deem A and B to be complements. If the interrelation is negative (arrow 2), we refer to A and B as substitutes [6]. Depending on the price difference between grade A and grade B, two products can be sold as complements or substitutes. The STSS should incorporate the interdependency between the apple grade. We call it intra-apple pricing effect.

On the other hand, obviously, apple has the substitutes such as peer and orange. However, the STSS does not incorporate all substitutes. Only other types of apples are include the modeling such as Huzi, Hongok, Hongru, and so on. We call it inter-apple pricing effect. Substitutes effect among apple types will be appeared by the price differences among them. So, we get

$$q_t = a_t - c_1 \sinh \left(\frac{c_2 (p_{t-1} - p_{t-1}' - (p_{t-2} - p_{t-2}'))}{(p_{t-2} - p_{t-2}')} \right) - c_3 \sinh \left(\frac{c_4 (p_{t-1} - p_{t-1}'' - (p_{t-2} - p_{t-2}''))}{(p_{t-2} - p_{t-2}'')} \right) \quad (4)$$

where $pt-1'$ = wholesale average price for intra-apple type at period t-1, $pt-2'$ = wholesale average price for intra-apple type at period t-2, $pt-1''$ = wholesale average price for inter-apple type at period t-1, $pt-2''$ = wholesale average price for inter-apple type at period t-2 and $a, c1, c2, c3,$ and $c4$ = parameters to be estimated.

2.4 Price Expectation Effects

In 1784, Adam Smith said that consumers have the tendency of postponement of purchase in anticipation of future lower price. Kucher(1985) calls it price expectation phenomenon[7]. Yoo(1991) was the first to empirically examine the expectation effect and its pricing implications for consumer durables[5]. Lee(2001) explored the price expectation phenomenon for 40 fashion brands[8].

In this regard, we can assume that a part of farmers have the tendency of postponement of shipment from warehouse or cold storage facilities to market in anticipation of future higher wholesale price. The STSS reflects this phenomenon as follows;

$$qt = a - b(Pt-1 + \beta t((pt-1-pt-2)/\tau) - c1 \sinh(c2(pt-1 - pt-1' - (pt-2 - pt-2')) / (pt-2 - pt-2')) - c3 \sinh(c4(pt-1 - pt-1'' - (pt-2 - pt-2'')) / (pt-2 - pt-2'')) \quad (5)$$

where βt = price expectation effect coefficient at period t and τ = discount rate.

2.5 Events Effect

On Chusok (Korean Thanksgiving Day) and the Lunar New Year's Day (Chunjie in China), customers feel different about brands, promotion, and product. Especially, when people prepare foods for ancestral rites table, the price elasticity for the foods is not sensitive. So, the price for apples was high at national holidays and the demand for apples was also high.

If the STSS is based on the sales DBs of the year and make just one price response function, we are likely to neglect the different aspects of dull and active seasons, promotion and non-promotion seasons, normal and abnormal seasons, and high-demand and low demand seasons. To reflect differences in customers' responses to price and to capture the abnormal demand and supply increase at national holiday seasons, the researcher adds the event coefficients into STSS.

$$qt = a - b(Pt-1 + \beta t((pt-1-pt-2)/\tau) - c1 \sinh(c2(pt-1 - pt-1' - (pt-2 - pt-2')) / (pt-2 - pt-2')) - c3 \sinh(c4(pt-1 - pt-1'' - (pt-2 - pt-2'')) / (pt-2 - pt-2'')) + \delta t Eventt \quad (6)$$

where δt = event effect coefficient at period t if the period t is holidays, $Eventt = 1$, otherwise = 0.

2.6 Demand Curve

The relationship between sales quantity and resulting price levels is called demand curve. The STSS needs to calibrate the demand curve to predict the price level for each period t until shipment by date. The shipment by date means the last day can be ship.

2.7 Shipment Timing Life Cycle

The period from the apples' first launch to the market until its final withdrawal at wholesales market at a given year is called 'apples' shipment Timing life cycle'. The period splits up in four phases; introduction, growth, maturity and decline phases. At the introduction phase, the shipment volume is small and the wholesale price is relatively high. At the growth phase, the shipment volume and the wholesale price increase in comparison to the introduction phase. The maturity phase shows the highest shipment volume and price. At the decline phase, the shipment volume and price decreases. According to the sub-period, the farmers have to control his shipment timing to wholesales market.

2.8 Cost Function and Profit Estimation

The cost function traces the changes in cost by apple unit when the storage time goes by. During storage in warehouse, the various costs happened such as storage cost, physical loss, down-grading, logistic cost, and price decrease possibility. The system calibrates the total cost at period t . Using formula 6 and 8, the system calibrates the expected income. Therefore, the system can estimate the profit resulting from the difference between the expected income and cost at every period t until $t+n$, shipment by date, as shown at Fig. 7. The researcher hopes that apple farmers and apple holders at storage can access to the application, where they find the profit and the expected income at period t from present time to shipment by date.

3 Conclusion

The purpose of this research is to develop a comprehensive shipment timing support system that applies to farmers using to decide shipment or not at a given time. The Shipment Timing Support System (STSS) consists of the dynamic price response function for each product, the price expectation effect, product line effect, substitutes price effect, events effect, the demand curve for apples and cost function. Based on STSS, the farmers can get opportunities to increase their profit.

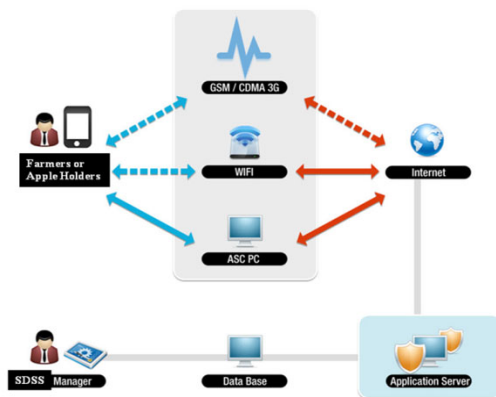


Fig. 6. Potential Network System for STSS

The fluctuation of the expected sales revenue for every product they store in warehouse will be provided by the application on smart phone as Fig. 7. On the UI, farmers also can find the timing for shipment. The STSS is evolutionary. Future version of STSS will include wider substitutes' price response to STSS[9]. And, the demand curve should incorporate the substitutes demand curve.

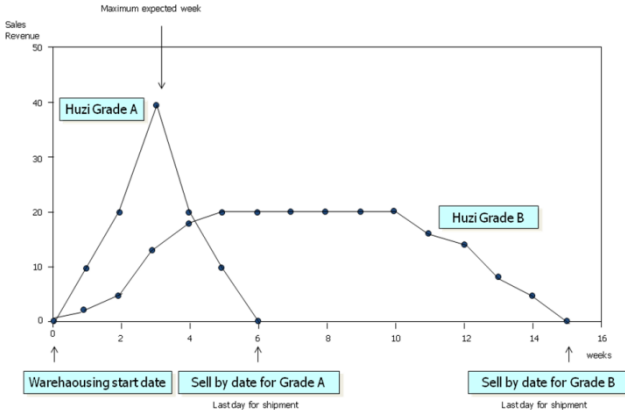


Fig. 7. Informations appeared on UI at STSS application

Acknowledgements. This work (Grants No.311024-3) was supported by a grant from the Ministry for Food, Agriculture, Forestry and Fisheries in 2011.

References

1. Simon, H.: Price Management. North-Holland, New York (1989)
2. Seo, S., Kim, T.: Operating Statuses of Cold Storage Facilities and Factors of Storage and Sales Decision-Making in Stored Apples. Korean Journal of Agricultural Management and Policy 36(1), 179–197 (2009)
3. Seo, S., Kim, T.: Stochastic Dominance of Expected Income by Sales Timing for Stored Apples in Korea. Journal of Korean Agricultural Economic Association 50(1), 85–101 (2009)
4. Kwak, Y., Hong, J., Lee, Y.: Strategic Pricing. Benet (2007)
5. Yoo, P.H.: Pricing Policy. Pakyoungsa, Seoul (1991)
6. Dolan, R., Simon, H.: Power Pricing. Free Press (1997)
7. Kucher: Scannerdaten und Preissensitivitat bei Konsumgutern. Gabler-Verlag, Wiesbaden (1985)
8. Lee, Y.: A Study on the Price Expectation Effect for Apparel Product based on Dynamic Price Response Function. Ph.D. dissertation, Dept. of Fashion, Sungkyunkwan Univ., Korea (2001)
9. Nagle, T., Holden: The Strategy and Tactics of Pricing. Englewood Cliffs (1995)

Heuristic for Paper Roll Selection for Corrugation Process Simulation

Chantana Chantrapornchai and T. Sathapanawat

Department of Computing, Faculty of Science, Silpakorn University
Thailand, 73000
ctana@su.ac.th

Abstract. In this research, we develop a software for production planning for the corrugation process. The software particularly creates a plan for paper roll selection for the mill stands. The plan is used by an employee to pick up paper rolls from the inventory and put on the stand. The main goal of the plan is to optimize the waste of the paper for each roll. There exists supplementary goals such as to balance paper roll distributions in all the mill stands. In this paper, such a problem is described. Then the greedy algorithm for roll selection is developed. The algorithm is integrated in the software to create a plan for the whole application.

Keywords: Corrugation Process, Planning, Paper roll selection.

1 Introduction

Planning and scheduling in a factory is considered as an NP-hard problem. Mathematic models are often needed to find an optimal solution. Many researches are focused on heuristics to find an approximate solution.

In this paper, we study a production planning problem for a corrugation process. In the paper factory, the mill stand is fed with the sequence of the paper rolls for corrugation. For a large factory, there are several mill stands and several workers. The worker needs to feed the roll to the mill stands. The sequence of paper rolls to be fed needs to be determined in advance. A planning algorithm which yields the sequence of roll feeding is studied. Our plan generates a sequence of feeding rolls to each mill stand under the constraints while maximizing the production goal. The problem is very similar to the traditional subset sum problem. Our algorithm is based on the greedy style which can vary a heuristic function. The algorithm will be integrated in the corrugation software which collaborates with existing database to determine paper rolls in stock and help inventory planning.

Several works have been done in cutting stock, scheduling and production planning. Most of them are solved using heuristic or linear programming approach. Nevertheless, very few of these has focused on the corrugation process. Also, existing software for corrugation process or production planning is commercial and very expensive.

For example, the worked by Dyckhoff (1981) presented the use of linear programming to solve 1-D cutting stock which is NP-complete problem. Afshar (2008) also used linear programming model to solve 1-D cutting stock with trim loss. Renio et.al. (2006) has given a good literature review on 1-D cutting stock problem.

For the scheduling problem, Pandelis G. Ipsilandis (2007) uses MOP in scheduling in project management. His model attempts to minimize the overall time, idle time as well as cost. Sang M. Lee et. al. (2003) used linear programming to solve job-scheduling and compared with the genetic algorithm. Chan et. al. (1998) proposed the machine scheduling using linear programming and compared the approach with list scheduling.

The closest work in the paper mill is the worked by Robert Hasseler (1980,1981,1990) which attempted to solve the trim loss for the corrugation process. He modeled using 0-1 linear programming. The trim loss problem is not considered in the original scheduling model. However, he did not consider the amount of butt rolls in the inventory and the selection of the paper roll to balance the mill feeding.

Several heuristics have been studied for scheduling problem. Kops et.al. (1994) presented a scheduling algorithm for single operation job flow. The approach used both linear programming and dynamic programming and considered the due date and minimizing the work in process. Haessler (1971) presented a heuristic to roll scheduling for cutting to minimize the setup cost, trim loss and total time based on the random solutions. Cruz et.al. (2002) proposed the technique for corrugating for corrugate sheet. Darley et.al.(2008) used agents for corrugate boxes in the factory. The agent analyzed the proper process to reduce inventory stock and considered the due date.

Many heuristics on the similar problem such as cutting stock problem are such as Shen et.al. (2007) which used particle swarm. Chiong and Beng (2007) compared genetic algorithm and evolutionary algorithm. Many commercial software exists for cutting stock problem. The closet ones are from SolarSoft (2008) which is for corrugate packing and from Mccullough (2007) from CTI company which presented a DSS system for a corrugating box for the whole production cycle.

2 Problem Backgrounds

In a corrugation box factory, corrugating is the most important process to save the cost. Normally, there are two types of corrugating boxes: Single wall and double wall corrugated boards as shown in Figure 1. In order to create the single wall board, 3 layers are needed while for the double wall board, 5 layers are needed.

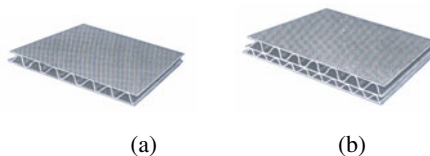


Fig. 1. Types of corrugating boards (a) Single wall (b) Double wall

In details the process will take the paper liner together with the paper medium and put together with the starch to become a piece of boards. Each paper is from the inventory which is in the roll as shown in Figure 2 (a). Each roll has a length of 4,000-6,000 meters. The unused roll is called full rolls while the already used one is called, butt rolls.



Fig. 2. (a) Paper rolls (b) Mill stands

Each paper roll is put on the mill roll stand in Figure 2(b). For each machine, there is two paper roll heads. The other one is the spare roll using when the first roll is run out. Thus, while the first roll is feeding, the worker needs to prepare the paper in the other roll and preparing to lengthen the paper.

Each paper has attributes such as the paper grade which implies the paper weight in grams and the width the paper which implies by the arrow in Figure 2(a). To calculate the paper size, the dimension of the paper box is used together with the total boxed needed plus the size to spare.

Thus, in our work, to select a proper roll, we consider the following (in the order of preference).

1. Select the roll that can be put on the mill roll stand in reality. That is the following cases:
 - The length of paper in the roll must be long enough so that the employee can setup the paper on the machine.
 - The roll that is selected on the i^{th} mill stand at time step a cannot be in the mill roll stand j at time b where $b-a \leq 5$
2. Total waste should be minimized. In the factory, the roll with less than 300 meters is a waste since it cannot be used on the mill stand anymore.
3. The butt rolls should be minimized. We will attempt to minimize the use of the new rolls.
4. To feed the rolls, we consider the small rolls (the rolls with the less paper, more than 300 meters but not more than 1,000 meters) in the first priority.
5. For the rolls with the attributes, we will select the older rolls. We will attempt to use the old one in the stock first to prevent the inventory aging.
6. For the worker on each mill stand, the work on roll changing should be balanced. If the workers are assigned with the small rolls, he will need to change the rolls very often.

3 Heuristic

Assume that at time step i at mill roll stand machine j there is a need for a paper as follows:

- paper width is w_i
- paper grade is g_i
- length is l_{ij}

There are the following constants that are configured.

- c_1 is the minimum length that can be on the mill stand machine e.g. 300 meters.
- c_2 is the length that can be wasted on the roll eg. 10 meters
- c_3 is the threshold to determine big roll or small roll eg. The paper roll that has length 1,500 meters or more is considered a big roll.
- c_4 is the number of jobs to be apart when the same roll is used again in the different mill stand e.g. there should be at least 5 jobs apart to reuse the same roll in another mill so that the employee can have sufficient time to move the roll.
- s_1 - s_5 is the total rolls accumulated at each mill stand for the whole plan, assuming there are 5 mill stands.

To search for a specific roll at time i mill stand j , we do the following

1. Search for roll R_n that is in the ready for use state which does not have a status marked as returne, or bad quality or in the c_4 job before at other mill stand.
2. From 1), search for the rolls that has the length exactly l_{ij} .
3. From 2), if not found, we search for the roll that has the total roll length L_n where $L_n - l_{ij}$ is smallest (or optimal fit) and $L_n - l_{ij}$ is greater than c_2
4. Alternative from 3), we try to find the roll that maximizes the value $L_n - l_{ij}$ (worst fit) so that the remainder can potentially be used.
5. Alternative from 4), check the previous roll if it is "big" roll or "small" roll using c_3 by trying to select the big roll alternates with the small roll until the whole required length l_{ij} is obtained while maintaining $c_2 < \text{waste} < c_1$ and minimizing waste.
6. For all the above cases, if there are more than 1 eligible rolls with the same specification, we select the roll that is the oldest by the manufacture date.
7. For the case that is no eligible roll, the notification is sent to the manager that the certain paper roll is out of stock so that the ordering can be processed.

We run the above algorithm for each time sequence and each mill stand. However in the above cases we have not consider to balance the load for each mill stand yet. This can be augmented by extra variables which is used to accumulate the number of changed rolls in each mill stand. The heuristic is attempt to select the best fit roll, then worst fit and then composite rolls. From this algorithm, the good rolls will be selected

in the beginning. In the later mill stand in the later sequence, which may require the same kind of paper, will get the smaller rolls. Then the employees at mill stand 3,4,5 on the later job, may be complained due to the changing of the rolls.

Thus, we have a little modification that uses the number of rolls at each mill stand to balance the roll selection. Then, the number of setup for each mill stand can be balanced.

4 Example

Figure 3 shows the schedule results. The results are read as follows: Column “No” shows the job number. Thus, the same row implies the same job. Column “Width” is the width of the roll. 5 mill stands will require the same width since they process the same job. Column “Grade#1” shows the paper grades needed for mill stand #1. Column “len#1” is the length required on mill stand #2 and so son. The empty entry in column “Grade” means there is no feeding at that sequence and the job uses the same paper grade as in the above row. For example, job no.3 at grade#3 “CA125” uses the same paper until job no.6. The whole length is 3,993 meters.

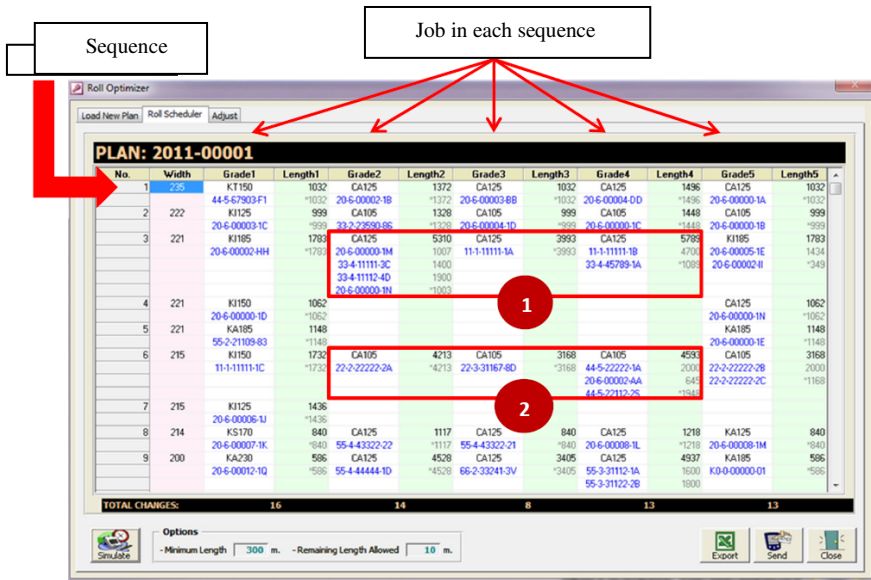


Fig 3. Example of the plan generation

From Figure 3, it is seen that at point 1 and point 2 at time step 6, the mill stand 2,3 and 4 need the same type of paper. Then we check that which mill stand among these has the most accumulated load. Then mill stand 2 will get the big roll first where the mill stand 4 gets the small rolls. This way, the number of setup will be more balanced.

5 Conclusion

We develop a heuristic for a paper roll feeding for corrugation process. The heuristic takes advantages of best fit and/or worst fit strategies. It creates a production plan for feeding a sequence of paper rolls for each stand mill. The algorithm considers many criteria in priority. The retrieved plan will help determine the location of paper rolls to be fed. Integrating with existing database, this will be advantages to the paper roll inventory planning.

Acknowledgement. We would like to thank National Research Council of Thailand for the funding support for this work.

References

1. Affshar, A., Amiri, H., Eshtelardian, E.: An improved linear programming model for one dimensional cutting stock problem. In: 1st Intl. Conf. on Construction in Developing Countries (ICCIDE-1), Pakistan (August 2008)
2. Chan, L.M.A., Muriel, A., Simchi-Levi, D.: Parallel Machine Scheduling, Linear Programming, and Parameter List Scheduling Heuristics. *Operations Research* 46(5), 729–741 (1998)
3. Chiong, R., Beng, O.: A comparison between genetic algorithms and evolutionary programming based on cutting stock problem. *Engineering Letter* 14(1) (2007) (Advance online publication February 12, 2007)
4. Cui, Y., Yang, Y.: A heuristic for one-dimensional cutting stock problem with usable leftover. *European Journal of Operation Research* 204(2010), 245–250 (2010)
5. Dyckhoff, H.: A new linear programming approach to the cutting stock problem. *Operations Research* 29, 1092–1104 (1981)
6. Santa Cruz, C.D., Sandstrom, W.R.: Method for bending a corrugated sheet. United States Patent 6354130 (2002)
7. Darley, V., von Tessin, P., Sandler, D.: An Agent-Based Model of a Corrugated Box Factory: Tradeoff between Finished-GoodStock and On-Time-in-Full Delivery, <http://www.santafe.edu/~vince/pub/CorrugatedBoxFactory.pdf> (accessed February 20, 2008)
8. Goldratt, E.M., Cox, J.: *The Goal*, 3rd Revised Edition. North River Press (2004)
9. Haessler, R.W.: A heuristic programming solution to a nonlinear cutting stock problem. *Management Science* 17, 793–802 (1971)
10. Haessler, R.W.: Production Planning and Scheduling for an Integrated Container Company. *Automatica* 21(4), 445–452 (1985)
11. Haessler, R.W., Brian Talbot, F.: A 0-1 Model For Solving The Corrugator Trim Problem, Working Paper No. 205, Division of Research, Graduate School of Business Administration. The University of Michigan (February 1980)
12. Haessler, R.W.: Short Term Planning and Scheduling in a Paper Mill, Working Paper No. 268, Division of Research, Graduate School of Business Administration. The University of Michigan (May 1981)
13. Haessler, R.W., Sweeny, P.E.: Evaluating the Impact of Small Changes in Ordered Widths on Trim Yields, Working Paper No. 640, Division of Research, Graduate School of Business Administration. The University of Michigan (July 1990)

14. Ipsilandis, P.G.: Multiobjective Linear Programming Model for Scheduling Linear Repetitive Projects. *Journal of Construction Engineering and Management* 133(6), 417–424 (2007), doi:10.1061/(ASCE)0733-9364(2007)133:6(417)
15. Kops, L., Natarajan, S.: Time partitioning based on the job-flow schedule — A new approach to optimum allocation of jobs on machine tools. *International Journal of Advanced Manufacturing Technology* 9(3), 204–210 (1994)
16. Lee, S.M., Asllani, A.A.: Job scheduling with dual criteria and sequence-dependent setups: mathematical versus genetic programming. *Omega* 32(2), 145–153 (2004)
17. Mccullough, J.: Improved Planning Results in Success, *Softwar Technology Case Studies*. Corrugated Technologies Inc. (January-February 2007)
18. Pinedo, M.L.: *Scheduling: Theory, Algorithms, and Systems*, 3rd edn. Springer, Heidelberg (2008)
19. Renildo, G., et al.: Linear Programming Models for the One-dimensional Cutting Stock Problem. INPE ePrint:sid.inpe.br/ePrint@80/2006/12.18.16.52 v1 2006-12-19 (2006)
20. Shen, X., Li, Y., Yang, J., Yu, L.: A Heuristic Particle Swarm Optimization for Cutting Stock Problem Based on Cutting Pattern. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloat, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4490, pp. 1175–1178. Springer, Heidelberg (2007)
21. SolarSoft Co. Ltd.,
http://www.solarsoft.com/industries/packaging/corrugated_packaging.htm

A Spatial Morphology of the English Verb Structure of Simple Sentences Aids Appropriate Cognitive Structures Develop and Offers EFL/ESL Potential as a Mobile App

Robert C. Meurant

Director, Institute of Traditional Studies, Seojeong College University, Yongam-Ri 681-1,
Eunhyun-Myeon, Yangju-Si, Gyeonggi-Do, Seoul, Korea 482-863

rmeurant@me.com

<http://web.me.com/rmeurant/INSTITUTE/HOME.html>

Abstract. For ubiquitous learning, appropriate learning objects need to be accessible at the right time and place. EFL/ESL learners face confusion in their use of appropriate English verb tenses and forms. The best place to locate such learning objects is in the learner's awareness, so that they may creatively imagine the necessary schema, navigate through it, and locate the correct formal expression of what they intend. I advance a spatial morphology of the English verb forms and tenses applied to generate simple sentences. This is designed to enable appropriate cognitive structures to form and stabilize in the language learner's mind while allowing ready on-screen implementation as an interactive mobile application for smartphones and tablets, head up displays, and online grammatical resources such as webpages. The primary intended realization is thus as an *aide-mémoire* in the learner's mind, and as an App for the iPad.

Keywords: EFL/ESL, English verb structure, simple sentences, mobile app, learning objects, spatial morphology, cognitive structure, iPad, iOS, CALL.

1 Introduction

As Fiaidhi observes [1], in the rush towards ubiquitous learning, appropriate learning objects need to be made accessible to learners at the right time and in the right place. EFL/ESL students understandably face confusion in their recall, choice and use of English verb tenses and forms that are deemed appropriate to an intended context. Of course, the best place to locate such learning objects is, through assimilation, in the learner's own awareness, so that they can creatively imagine the necessary schema, navigate through it, and locate the correct formal expression of what is intended. Providing an adequate spatial morphology of these tenses and forms offers the potential of more effective learning through cognitive schema and interactive apps.

In this paper, I advance a spatial morphology of the English verb forms and tenses applied to generate simple sentences, which is designed to enable appropriate cognitive structures to be formed and stabilized in the language learner's mind. The visual spatial structure is designed to allow ready implementation in interactive digital resources (in mainly landscape mode), for use in software driving on-screen displays.

These include mobile apps for devices such as smart phones (which students are likely to have with them), and more particularly in tablets such as the iPad (which will increasingly be used by students); head up displays (e.g. by computer gamers, Sloodle users, pilots, vehicular drivers etc.); and for electronic dictionaries and online resources such as online dictionaries and other grammatical resource webpages. The primary intended realization of this is two-fold: firstly as an *aide-mémoire* in the learner's mind; and secondly as an App for the iPad. I envisage that having students use the App would significantly assist their learning of correct cognitive structures.

2 Structure and Use of the English Verb App

2.1 The Spatial Morphology of the English Verb App

The spatial morphology (used not in a linguistic sense, but in the traditional sense of formal organization) draws its inspiration from certain configurations common in Sacred Geometry, Art and Architecture that have traditionally served to enable ready learning and contemplation of highly complex cosmologies and the realities that they are deemed to mediate. The primary geometrical form is of the aedicule, represented as a flat vertical rectangular surface in the normal field of vision, rather than as a horizontal plan. At the center is situated the Infinitive form of the verb, the basic and most essential form that the verb takes. The surface is developed as a four-fold inner zone, which for a given verb displays its four aspects, sentences that accord with those aspects, their construction, and use. A peripheral border displays other grammatical categories. For the Finite verbs and forms, the bottom left inner quadrant is the starting point for the four fundamental grammatical Aspects, and represents the Simple verb. Extending the Simple form to the right gives the Progressive/Continuous form of the verb at bottom right; while (as it were an octave of) the Simple, directly above, gives the Perfect form at top left. Finally, extending that Perfect form to the right, while simultaneously providing (a hovering octave of) the Progressive directly above, gives the Perfect Progressive/Continuous at top right.

The border around the four edges accommodates the other major grammatical categories, and comprises interactive buttons, tabs, and where needed drop-down menus. The important TENSE-ASPECT-MODE dimensions are thus provided vertically, as in Fig. 1. The base form of the verb (the bare Infinitive) is entered or selected from a drop-down menu from the right-hand major top field. Side options of the top and bottom borders provide for grammatical categories of QUALITY (i.e. Finiteness), NUMBER, VOICE and POLARITY. The top right Number field works in conjunction with the PERSON and GENDER categories that are provided in the right side border. MODE, in the bottom border, is differentiated into major fields of MOOD and MODALITY. Contextual fields further expand these options in the left side border. For the purposes of this paper, verbs are presumed to be either Intransitive or Optionally Transitive, and sentence objects are not required nor displayed. (The left border could instead display TRANSITIVITY fields accommodating Obligatory/Optional Transitive/Intransitive, and interact with verb PROCESS Dynamic/Stative, and verb VALENCY 1/2/3 could be included; see [2]).

2.2 Use of the English Verb App

The ESL/EFL learner firstly enters a base verb (the bare Infinitive form of the verb) at the top right, either by selecting from a drop-down menu, or by typing (with intelligent word recognition and completion); this could display in the center as in Fig. 2B. Default settings for the grammatical categories would accord with the most common states for English of the target learner level, but could be reset in a Settings menu, which also accommodates choice of UK, US or other dialect. The user then selects options from the border fields in accord with the desired use. The minor side fields on the top and bottom borders represent binary choices, and are exclusive “or” choices except for NUMBER, which can be deselected, have either option selected, or both. The top border QUALITY (Finiteness) provides exclusive choice of Finite or Non-finite Verb form; at bottom, VOICE allows for exclusive Active or Passive choice, while POLARITY allows for exclusive Affirmative or Negative choice.

Major top, bottom and right border fields allow for more than two options. TENSE is contextual, and when Finite Quality is selected, provides for an exclusive Past, Present or Future choice, provided the MODALITY choice is non-modal. When the QUALITY is Non-finite, the options will depend upon other settings e.g. whether the VOICE is Active or Passive. Along the bottom border, MOOD at left provides for exclusive Declarative (Indicative) “declare”, Interrogative “question”, Imperative “direct” or Subjunctive “subjunct” choices. The contextual menu along the left side border may be activated by these choices, e.g. when the Interrogative Question is selected, the left border could show Wh- question words who/what/where etc. as prompts and/or could toggle with Polar question words such as do/did/is/are/will etc. The bottom border at right provides for the exclusive selection of MODALITY: Non-modal, Modal, Semi-modal, or Other modal. When Modal is selected, the contextual menu of the left side border displays the Modals: may/might/can/could etc.

The right side border differs in that the choices need not be exclusive but as for NUMBER may be (partially or totally) inclusive and are intended to function in conjunction with the Number selection (at right of the top border). So PERSON can be selected as none, any, some or all of First/Second/Third; while GENDER can be selected as none, any, some or all of Male/Female/Neutral - in English it is only relevant if Third Person Singular has been selected. These provisions of exclusivity allows the display of just one PERSON, or of a limited, or of a full conjugation.

The App then displays in the inner quadrant the four Aspects of the specific verb base and chosen grammatical categories, showing the four simple sentences, sentence constructions, tenses, and typical uses (from Swan [3]). The same Verb form thus assumes the same relative position in each inner rectangle, so that for any given TENSE, the relation between Simple, Continuous, Perfect, and Perfect Continuous aspects is readily apprehended. Ideally, the learner internalizes the language structure.

2.3 Using Finger Gestures in the English Verb App

When the FINITE quality is chosen, single tapping any one ASPECT (of an inner quadrant) zooms that quadrant to fill the entire inner space, and would allow more detailed information to be displayed. A full conjugation could be displayed, as in Fig.

2c. Single tapping a zoomed interior collapses it to return to the four-quadrant view. All finger gestures would be in accord with Apple's iOS Human Interface Guidelines.

Horizontally swiping the interior permits cycling through the TENSEs, so a right swipe regresses the Tense being displayed, while a left swipe advances the Tense e.g. from Past to Present, or Present to Future. The Settings could allow the horizontal swipe to also include the Infinitive, cycling through Past-Present-Future-Infinitive.

When just one NUMBER and PERSON has been selected, vertical swiping of the interior zone allows cycling through the six Number/Person combinations 1PS-2PS-3PS-1PP-1PP-3PP (First Person Singular - Second Person Singular - Third Person Singular - First Person Plural - Second Person Plural - Third Person Plural). Swiping the interior quadrants upwards will then advance the Number/Person displayed e.g. 2PS > 3PS; swiping it downwards will regress the display e.g. 2PP > 1PP.

3 Merits and Limitations of the English Verb App

This App encourages playful exploration of the effects of varying the values of different grammatical categories and thus allows a deepening appreciation of the structure of the English language, albeit at a relatively simple level.

Further, correspondences between verb forms for a particular tense can be comprehended in relation to the other tenses, for example in considering the relationships revealed by a horizontal left swipe sequence of Past Progressive below to Past Perfect Progressive above, Present Progressive below to Present Perfect Progressive above, and Future Progressive below to Future Perfect Progressive above. It is then apparent that the same Verb form assumes the same relative position in each configuration, so that for example the Perfect Past is in top left quadrant of the Past tense, the Perfect Present is in top left quadrant of the Present tense; and the Perfect Future is in top left quadrant of the Future tense. This structure should assist learning.

The limitations that are met are primarily those of the structural complexity of language in general and of the English language in particular (expert Grammarians are not infrequently at odds over issues of linguistic structure); and these must be set against the honest and understandable limitations of the second language learner. Thus a satisfactory compromise needs to be made between a simplicity that can be comprehended, and a complexity that does justice to the target language. To that end, the structural morphology presented does not address sentence objects (though these could be readily incorporated), and is restricted to simple sentences (it does not deal with compound and complex sentences). The treatments of passive and of non-finite forms have been greatly simplified; and imperative and subjunctive moods as well as semi-modal and other modalities, though provided for, have not been structured.

The App is designed to fit in broad agreement with the Apple iOS Human Interface Guidelines [4]: the display resolution accords with the 960x640 pxl iPhone screen, and fits comfortably on the 1024x768 pxl iPad screen. Border widths of 50 pxl at top/bottom and 48 pxl left/right mean both dimensions of the 864x540 pxl interior are rich in factors, which allows convenient subdivision into equal columns and rows ($864 = 2^3 \cdot 3^3$, $540 = 2^2 \cdot 3^3 \cdot 5$), while minimum button sizes are adequate. (Color-coding of buttons and tabs is not shown in this black & white paper; images are much reduced).

P	QUALITY	TENSE			VERB				NUMBER	PC
MOOD and MODALITY - EXPANDED	PERFECT				PERFECT CONTINUOUS				PERSON	
	ASPECT									
	SIMPLE				CONTINUOUS				GENDER	
S	VOICE	MOOD			MODALITY				POLARITY	C
P	finite / non-finite	past / present / future			VERB				singular / plural	PC
- will vary according to mood and modality selection -	perfect <small>(typically: connection between events in time; completion of something by a particular time)</small>				perfect continuous <small>(typically: continuity up to a particular time)</small>				first / second / third	
	simple <small>(typically: events in time; permanent situations)</small>				continuous <small>(typically: events as going on or continuing, perhaps at or up to a particular time)</small>				male / female / neutral	
S	active / passive	declare/question/direct/subjunct			nonmodal/modal/semimodal/other				affirmative / negative	C
P	1 ∞	past	present	future	VERB				□ <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	PC
- will vary according to mood and modality selection -	QUALITY (FINITENESS)	TENSE			VERB				NUMBER	1st
	MOOD and MODALITY - EXPANDED	PERFECT perfect				PERFECT CONTINUOUS perfect continuous				PERSON
ASPECT										
	SIMPLE simple				CONTINUOUS continuous				GENDER	
S	VOICE	MOOD			MODALITY				POLARITY	C
	-	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+ -

Fig. 1. The basic structure showing a. the main grammatical categories; b. differentiation of the main grammatical categories into options; and c. various options as actually presented (with categories appended in grey text)

P	1	∞	past	present	future	VERB				<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	PC	
whose	_ have VERB-en <i>perfect</i>					_ have been VERB-ing <i>perfect continuous</i>					1st		
whom											2nd		
how											3rd		
which	_ VERB <i>simple</i>					_ be VERB-ing <i>continuous</i>					♂		
why											♀		
when											♂ ♀		
where													
what													
who													
S		—	declarative mood			non-modal				+	-	C	
P	1	∞	past	present	future	walk				<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	PC	
whose	<i>I have walked</i>					<i>I have been walking</i>					1st		
whom	_ have VERB-en <i>present perfect</i> <small>(typically: past action with some present connection)</small>					_ have been VERB-ing <i>present perfect continuous</i> <small>(typically: continuity up to the present)</small>					2nd		
how						(to) walk					3rd		
which	<i>I walk</i>					<i>I am walking</i>					♂		
why	_ VERB <i>present simple</i> <small>(typically: general time; permanent situations)</small>					_ be VERB-ing <i>present continuous</i> <small>(typically: actions continuing at the moment of speaking)</small>					♀		
when											♂ ♀		
where													
what													
who													
S		—	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	-	C
P	1	∞	past	present	future	walk				<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	PC	
whose	<i>I have been walking</i>					_ have been VERB-ing <i>present perfect continuous</i> <small>(typically: continuity up to the present)</small>					1st		
whom	<i>You have been walking</i>										2nd		
how	<i>He/She/It has been walking</i>										3rd		
which											♂		
why											♀		
when											♂ ♀		
where													
what													
who													
S		—	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	-	C

Fig. 2. Finite/Present/Active/Declarative/Non-modal/Affirmative showing: a. archetypal sentence construction; b. with Singular/First Person/ sentence construction (Infinitive at center could be tap-toggled on and off); c. full conjugation for all Numbers, Persons and Genders.

P	1	∞	past	present	future	walk	□	▣	▤	▥	PC			
whose	you had walked					you had been walking					1st			
whom	_ had VERB-en					_ had been VERB-ing					2nd			
how	past perfect					past perfect continuous					3rd			
which	(typically: action before a particular past time)					(typically: continuity up to a particular past time)								
why	you walked					you were walking					♂			
when	_ VERB-ed					_ been VERB-ing					♀			
where	past simple					past continuous					♂♀			
what	(typically: past events)					(typically: actions continuing at a particular past time)								
who	S	I	—	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	—	C
whose	P	1	∞	past	present	future	eat	□	▣	▤	▥	PC		
whose	they haven't eaten					they haven't been eating					1st			
whom	_ have not VERB-en					_ have not been VERB-ing					2nd			
how	present perfect					present perfect continuous					3rd			
which	(typically: past action with some present connection)					(typically: continuity up to the present)								
why	they don't eat					they aren't eating					♂			
when	_ don't VERB					_ be not VERB-ing					♀			
where	present simple					present continuous					♂♀			
what	(typically: general time; permanent situations)					(typically: actions continuing at the moment of speaking)								
who	S	I	—	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	—	C
whose	P	1	∞	past	present	future	walk	□	▣	▤	▥	PC		
whose	will he/she/it have walked?					will he/she/it have been walking?					1st			
whom	will _ have VERB-en?					will _ have been VERB-ing?					2nd			
how	future perfect					future perfect continuous					3rd			
which	(typically: action before a particular future time)					(typically: continuity up to a particular future time)								
why	will he/she/it walk?					will he/she/it be walking?					♂			
when	will _ VERB?					will _ be VERB-ing?					♀			
where	future simple					future continuous					♂♀			
what	(typically: information about the future; future events)					(typically: actions continuing at a particular future time)								
who	S	I	—	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	—	C

Fig. 3. a. Finite/Past/Singular/Second Person/Active/Declarative/Non-modal/Affirmative; b. Finite/Present/Plural/3P/Active/Declarative/Non-modal/Negative with different verb; and c. Finite/Future/Singular/3P/Active/Interrogative/Non-modal/Positive constructions.

P	1	∞	past	present	future	walk	□	□□	□□	PC			
whose	she has been walked					she has been being walked (?)					1st		
whom	_ have been VERB-en					_ have been being VERB-en					2nd		
how	present perfect passive <i>(typically; past action with some present connection)</i>					present perfect continuous passive <i>(typically; continuity up to the present)</i>					3rd		
which													
why	she is walked					she is being walked					♂		
when	_ be VERB-en					_ be being VERB-en					♀		
where	present simple passive <i>(typically; general time; permanent situations)</i>					present continuous passive <i>(typically; actions continuing at the moment of speaking)</i>					♂		
what													
who													
S		-	declare	question	direct	subjunct	non-modal	modal	semi-modal	other	+	-	C

Fig. 4. Finite/Singular/3P/Passive/Declarative/Non-modal/Affirmative: Present/Female. (Bracketed terms help identify the appropriate tense).

4 Conclusion

As I elsewhere detail, it is now apparent that the ICT revolution is radically impacting pedagogy [5]. In language learning, this is now noticeable in the cognitive schema that might be effectively utilized in conjunction with language learning strategies that are more suited to the digital age. Fiaidhi rightly maintains that the pervasive impact of the Internet, mobile apps, smartphones and now tablets means that learning objects are required that satisfy demand for ubiquitous learning at the right time and in the right place. The ways in which language is conceived, imagined, learnt, and utilized are changing. Envisaging and developing spatial morphologies that aid learning - via a dialectic between the space of the imagination and the digital realm of mobile apps - will I trust contribute to that exciting educational and technological evolution.

Acknowledgments. My thanks for intellectual inspiration are due to Daniel Howard of Howard Science Ltd, and to linguist Gavin Austin of the School of Behavioural, Cognitive and Social Sciences, University of New England (Australia).

References

1. Fiaidhi, J.: The Impact of Ubiquity on Learning Objects. Keynote, UCMA, Daejeon (2011)
2. Crowley, T., Lynch, J., Siegel, J., Piau, J.: The Design of Language: An Introduction to Descriptive Linguistics. Longman, Auckland (1995)
3. Swan, M.: Practical English Usage, 3rd edn. Oxford University Press, Oxford (2005)
4. Apple Inc.: iOS Human Interface Guidelines. Apple Inc., Cupertino (2011)
5. Meurant, R.C.: Applied Linguistics and the Convergence of Information Communication Technologies: Collected refereed papers 2006-2009. The Opoutere Press, Auckland (2010)

An Empirical Analysis of the Effects of IT Professionals' Emotional Dissonance on Creativity Revelation Processes and Individual Creativity

Seong Wook Chae¹, Young Wook Seo², and Kun Chang Lee^{3,*}

¹ Principal Researcher, National Information Society Agency, Republic of Korea

² Researcher, Software Engineering Center at NIPA, Republic of Korea

³ Professor, SKK Business School and Department of Interaction Science, Sungkyunkwan University, Seoul 110-745, Republic of Korea
{seongwookchae, seoyy123, kunchanglee}@gmail.com

Abstract. Employees in frontline service jobs are susceptible to emotional dissonance, or incongruence between one's experienced and displayed emotions. The purpose of this paper is to empirically investigate the effect of emotional dissonance and job stress on the individual creativity of frontline IT professionals through creativity revelation processes such as exploration and exploitation. To validate our hypotheses, we collected questionnaires from 447 frontline IT workers responsible for creative output by communicating with organizational counterparts, and analyzed the data using a structural equation model. The results show that emotional dissonance has a positive influence on exploitation, and job stress has a negative impact on exploration.

Keywords: Individual creativity, Emotional dissonance, Exploration, Exploitation, Job stress.

1 Introduction

Today's information technology workers are facing increased complexity of both technology and its creative use. They are required to have problem solving and customer service expertise in addition to technical skills [1, 2]. Some IT professionals are regularly expected to work with colleagues in IT and other areas of the organization, and are expected to conform to occupational or organizational norms regarding the display of emotion during these interactions. Frontline IT employees experiencing frequent face-to-face or voice-to-voice interactions with customers must understand the business and communicate effectively with organizational counterparts. However, employees in frontline service jobs are susceptible to emotional dissonance [3], which refers to incongruence between one's felt and displayed emotions. Emotional dissonance has generally been regarded as a negative influence on job satisfaction but a positive impact on work exhaustion, turnover rate, and stress. However, considering creativity revelation and stressful job situations, we have questions about how negative psychological factors affect the creativity of

* Corresponding author.

frontline IT workers. The purpose of this study is thus to examine the effect of frontline IT professionals' emotional dissonance and job stress on individual creativity through creativity revelation processes such as exploration and exploitation.

2 Literature Review

2.1 Emotional Dissonance and Job Stress

Emotional dissonance has been examined in the context of human service occupations such as flight attendant, bank teller, and bill collector. Employees in these jobs are often expected to present a desired emotional demeanor at work. Firms may explicitly mandate these expectations, termed display rules, to further organizational goals such as. To cope with the firm's display rules, employees can choose from two types of action. One is 'deep acting' and occurs when an employee tries to regulate felt emotions in order to comply with the display rules. The second is 'surface acting' and involves an employee faking his/her emotions by changing his/her outer demeanor to conform with the display rules while the inner feelings remain unchanged [4, 5].

Emotional dissonance has been conceptualized by researchers in a variety of ways. First, there is the point of view that emotional dissonance is an antecedent of emotional labor [4, 6], which is defined as "the effort, planning, and control needed to express organizationally desired emotion during interpersonal transactions" (Morris and Feldman, 1996, p. 987) [7]. Another point of view is that emotional dissonance is a contributing part or factor of emotional labor [8-10]. In other words, emotional dissonance is the action (or labor) itself that suppresses perceived emotions or expresses emotions that did not originally occur. Finally, emotional dissonance is regarded by some as an employee's state of psychological conflict as a result of his emotional labor [11, 12]. According to Morris and Feldman (1997), emotional dissonance relates positively to work exhaustion and negatively to job satisfaction [8]. Emotional dissonance creates employee strain, which may lead to lowered self-esteem, alienation from work, and depression [13].

Job stress is defined as an uncomfortable and undesirable feeling experienced by an individual "who is required to deviate from normal or self-desired functioning in the work place as the result of opportunities, constraints, or demands relating to potentially important work-related outcomes" (Parker and DeCotiis, 1983, p. 165) [14]. Job stress has been linked to a decrease in organizational effectiveness, individual performance, and individual health and illness.

2.2 Exploration and Exploitation

The exploration and exploitation concepts introduced by March (1991) are important mechanisms related to organizational growth and survival. Exploration is explained through the use of terms such as search, variation, risk-taking, experimentation, play, flexibility, discovery, and innovation, and exploitation is related to refinement, choice, production, efficiency, selection, implementation, and execution. Exploration involves pursuing change, taking risk, and pushing for experiments, while exploitation tends to

reduce change and focus on efficiency [15]. The two concepts require different structures, processes, strategies, capabilities, and cultures, and affect organizational achievement.

2.3 Creativity

Creativity is generally defined as the outcome of a new, useful idea and as a resolution to a problem. It is also referred to as a process of creating an idea or resolving a problem as well as an actual idea or resolution itself [16, 17]. Research on creativity started with the exploration of individual cognitive and personality traits. Guilford suggested that creativity is a continuous trait in all people and that individuals with recognized creative talent simply have “more of what all of us have” [18]. Since Guilford’s study, researchers have mainly focused on “individual” creativity, which is an intersection of individual domain-relevant skills, creativity-relevant skills, and motivation [19]. Creativity does not occur spontaneously or randomly but, instead, occurs along with the combination of knowledge, skill, and motivation that helps an individual generate new ideas. Creativity happens in the interaction between a person’s thoughts and the sociocultural context, and is not only inside people’s heads. It can be viewed as a systemic rather than individual phenomenon [20]. In this way, research on individual creativity has mostly centered on personal and contextual characteristics and their interactions [21].

3 Research Model and Hypotheses

We developed a research model (Fig. 1) that consists of five constructs: emotional dissonance, job stress, exploitation, exploration, and individual creativity.

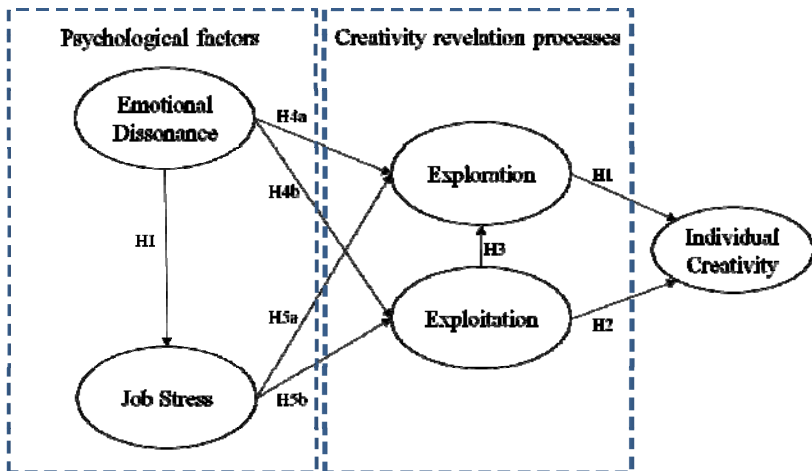


Fig. 1. Research model

3.1 Creativity Revelation Process and Individual Creativity

Exploration is an innovative activity in which new solutions are found and knowledge is added while exploitation refers to taking advantage of existing knowledge and resources and seeking more differentiated ones [15]. Approaches to exploration and exploitation have mostly been taken from the organizational study perspective, discussed in the process of producing creative or innovative outcomes [22-24]. Audia and Concalo (2007) divided creativity into divergent creativity and incremental creativity, explaining divergent creativity by linking it with exploration activity while elucidating incremental creativity by connecting it to exploitation activity [22]. Benner and Tushman (2002) divided innovation into exploratory innovation and exploitative innovation. Exploratory innovation is referred to as exploring a new competence and developing new technological trajectories other than the existing technological trajectories owned by an organization [23]. Therefore, exploration and exploitation activities may naturally be related to creative endeavors. We suggest that exploration and exploitation stimulate individual creativity and thus propose the following hypotheses.

Hypothesis 1: Exploration positively influences individual creativity.

Hypothesis 2: Exploitation positively influences individual creativity.

Hypothesis 3: Exploitation positively influences exploration.

3.2 Psychological Factors and the Creativity Revelation Process

In this study, emotional dissonance is regarded as an emotional labor factor. Hence, it is defined as the action (or labor) to suppress perceived emotions or express emotions that are different from those that originally occurred [8-10]. Frontline employees are required to follow organizational display rules for better organizational goal performance. In order to cope with the required display rules of the organization, employees can engage in surface acting or deep acting. Although an employee might have emotional dissonance through surface acting, he can eventually meet the organizational display rules during communicating with organizational counterparts. From that point of view, emotional dissonance is expected to have a positive impact on creativity revelation processes, and we propose the following hypotheses.

Hypothesis 4a: Emotional dissonance positively influences exploration.

Hypothesis 4b: Emotional dissonance positively influences exploitation.

However, emotional dissonance refers to incongruence between felt and displayed emotions, and may create employee strain which can then lead to lowered self-esteem, alienation from work, and depression [13]. Job stress is an uncomfortable and undesirable feeling resulting from work-related outcomes, and has been linked to failing individual health and illness, decreased individual performance, and decreased organizational effectiveness. Hence, we formed the following hypotheses.

Hypothesis 5a: Job stress negatively influences exploration.

Hypothesis 5b: Job stress negatively influences exploitation.

Hypothesis 6: Emotional dissonance positively influences job stress.

4 Empirical Analysis

4.1 Data Collection

The purpose of this study is to examine the effect of emotional dissonance and job stress on the creativity revelation process in the context of the information technology (IT) business environment. Data for this study were collected through a questionnaire administered to frontline IT professionals responsible for creative output through communication with organizational counterparts. In our analysis, we used 447 valid questionnaires. Respondents' major roles included IT consultant (25.1%), IT planning (22.1%), system analyst (13.0%), and requirement analyst (9.2%). About 82% of respondents were male, with 17.7% female, and 61.7% of respondents were in their thirties.

4.2 Measurement

The questionnaire we developed aimed to measure emotional dissonance, job stress, exploration, exploitation, and individual creativity. All measurement items on the questionnaire were measured on a seven-point Likert-type scale, with answers ranging from "strongly disagree" (1) to "strongly agree" (7). Most items on the survey were directly adopted from the existing literature and were previously validated by other researchers. In the case of individual creativity, survey items were adapted from Zhou and George [25]. Exploitation and exploration were measured using seven items adapted from Prieto et al. [26]. The items used to measure emotional dissonance, negative affectivity, and job stress were developed from previous studies by Rutner et al. [10] and Sosik and Godshalk [27]. In addition, we included negative affectivity, sex, and organizational tenure as control variables.

4.3 Reliability and Construct Validity

By referring to SmartPLS Version 2.0, partial least squares (PLS) was used for measurement validation and testing the structural model shown in Fig. 1. PLS places minimal restrictions on the sample size and residual distribution [28], so it is particularly useful in areas where there is weak theory and limited understanding of the relationships between variables [25].

We conducted reliability and validity analyses to examine if the questionnaire items matched our intent. The items were also tested for scale reliability. The Cronbach's alpha scores all exceeded 0.7, indicating high internal consistency. The convergent validity was first assessed by reviewing the t-test results for factor loading and was then assessed by examining the composite reliability and average variance extracted (AVE). Discriminant validity was assessed by examining the composite reliability and average variance extract (AVE; See Table 1).

Table 1. Reliability and convergent validity

Construct	Measurement Item	Factor loading	t-value	Composite reliability	Cronbach a	AVE
Individual Creativity (IC)	CR1	0.833	43.012	0.932	0.902	0.773
	CR2	0.891	81.505			
	CR3	0.895	84.947			
	CR4	0.897	79.524			
Exploration (ER)	ER1	0.861	58.788	0.908	0.847	0.766
	ER2	0.902	94.936			
	ER3	0.862	70.479			
Exploitation (ET)	ET1	0.819	45.698	0.878	0.813	0.643
	ET2	0.843	60.333			
	ET3	0.737	25.307			
	ET4	0.804	41.049			
Emotional Dissonance (ED)	ED1	0.845	46.547	0.954	0.945	0.723
	ED2	0.809	38.734			
	ED3	0.886	79.696			
	ED4	0.878	73.773			
	ED5	0.850	49.639			
	ED6	0.843	41.394			
	ED7	0.855	54.320			
	ED8	0.832	44.383			
Job Stress (JS)	ST1	0.821	41.797	0.935	0.914	0.743
	ST2	0.844	49.363			
	ST3	0.875	66.204			
	ST4	0.885	79.575			
	ST5	0.885	75.328			

For discriminant validity, a measure should more strongly correlate with all measures of the same construct than it does with any measures of the other constructs. For satisfactory discriminant validity, the average variance extracted (AVE) from the construct should be greater than the variance shared between the construct and other constructs in the model. Table 2 presents the correlation matrix with the correlations among the constructs and the square root of AVE on the diagonal. As shown in Table 2, the discriminant validity was at an acceptable level.

Table 2. Discriminant validity

Construct	ED	ER	ET	IC	NA	ST
ED	0.850					
ER	0.338	0.875				
ET	0.499	0.590	0.802			
IC	0.210	0.645	0.458	0.879		
NA	0.028	-0.055	-0.113	-0.052	0.896	
ST	0.082	-0.088	-0.069	0.009	0.638	0.862

※ Note: The diagonal element of correlation is the square root of AVE. To support discriminant validity, the diagonal element needs to be higher than the values of the off-diagonal area.

4.4 Results

The structural model can be calculated as shown in Fig. 2. First, as hypothesized, exploration ($b=0.552, p<0.001$) and exploitation ($b=0.133, p<0.01$) significantly influenced individual creativity, thus validating H1 and H2, respectively. Second, the coefficient between exploitation and exploration was significant ($b=0.561, p<0.001$), supporting H3. Next, emotional dissonance had a positive impact on exploitation ($b=0.501, p<0.001$), but not exploration ($b = 0.065, n/s$) or job stress ($b = 0.065, n/s$). These results support H4b, but not H4a or H6. Finally, job stress significantly influenced exploration ($b=-0.105, p<0.05$), but not exploitation ($b=-0.044, n/s$), thus validating H5a but not H5b, respectively.

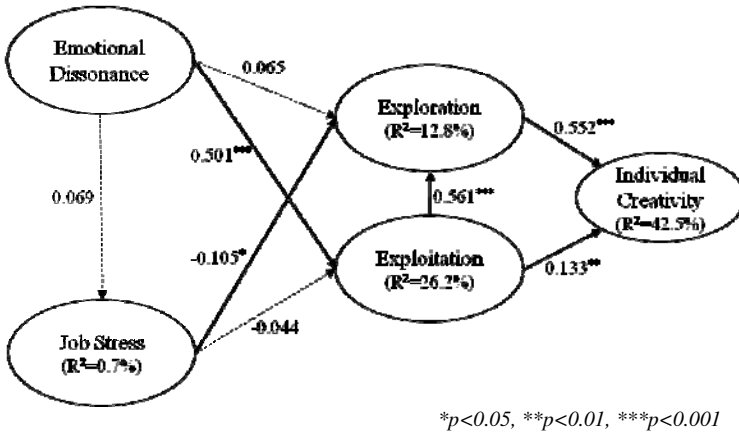


Fig. 2. Results

5 Discussion and Conclusions

We implemented an empirical test to investigate the impact of emotional dissonance and job stress on frontline IT professionals' individual creativity through exploitation and exploration. The results show that, first, emotional dissonance positively influences exploitation. Frontline IT workers, especially IT consultants, system analysts, and requirement analysts, must successfully communicate with customers in order to produce creative output. To perform their jobs effectively, they sometimes have to act as if they empathize with their business partner despite their actual lack of concern, or pretend to not be irritated with their organizational counterparts even when they may feel that way. Emotional dissonance also indicates that the IT worker is meeting the organizational display rules required by his or her job through surface acting. As a result, their creativity revelation activity will be successful. Second, exploitation fully mediates the effect of emotional dissonance on exploration. While emotional dissonance was found to significantly influence exploitation, it did not directly affect exploration. This study also revealed that emotional dissonance impacts

exploration only through exploitation. IT workers engage in surface acting, which leads to an uncomfortable mental status although exploration is a more difficult process than exploitation in that it requires something new and novel, while exploitation starts with existing knowledge. Third, job stress has a negative impact on exploration but not on exploitation. This study showed that IT employees could conduct exploitation activities regardless of their job stress, but the findings related to exploration were different. Exploitation using existing knowledge is a relatively easier process than exploration, which involves making something new. In other words, if we expect to obtain creative output through exploration activities, IT professionals' job stress should be decreased.

Future research on the relationship between emotional dissonance and creativity revelation processes is needed. Although this study showed that IT professionals' emotional dissonance could facilitate their creativity revelation processes, it is not certain that this creativity will last in the long term because they must devote mental resources toward coping with that dissonance.

Acknowledgment. This research was supported by the World Class University (WCU) program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology, Republic of Korea (Grant No. R31-2008-000-10062-0).

References

1. Pawlowski, S.D., Robey, D.: Bridging User Organizations: Knowledge Brokering and the Work of Information Technology Professionals. *MIS Quarterly* 28, 645–672 (2004)
2. Thibodeau, P.: Diverse Skills Can Help IT Workers Avoid Offshore Ax. *Computerworld* 38, 16–17 (2004)
3. Phillips, B., Tan, T.T.W., Julian, C.: The theoretical underpinnings of emotional dissonance: a framework and analysis of propositions. *Journal of Services Marketing* 20, 471–478 (2007)
4. Zapf, D., Vogt, C., Seifert, C., Mertini, H., Isic, A.: Emotion Work as a Source of Stress: The Concept and Development of an Instrument. *European Journal of Work and Organizational Psychology* 8, 371–400 (1999)
5. Hochschild, A.: *The Managed Heart*. University of California Press, Berkeley (1983)
6. Brotheridge, C.M., Grandey, A.A.: Emotional Labor and Burnout: Comparing Two Perspectives of “People Work”. *Journal of Vocational Behavior* 60, 17–39 (2002)
7. Morris, J.A., Feldman, D.C.: The dimensions, antecedents, and consequences of emotional labor. *Academy of Management Review* 21, 986–1010 (1996)
8. Morris, J.A., Feldman, D.C.: Managing emotions in the workplace. *Journal of Managerial Issues* 9, 257–274 (1997)
9. Schaubroeck, J., Jones, J.R.: Antecedents of workplace emotional labor dimensions and moderators of their effects on physical symptoms. *Journal of Organizational Behavior* 21, 163–183 (2000)
10. Rutner, P.S., Hardgrave, B.C., McKnight, D.H.: Emotional Dissonance and the Information Technology Professional. *MIS Quarterly* 32, 635–652 (2008)

11. Lewig, K.A., Dollard, M.F.: Emotional dissonance, emotional exhaustion and job satisfaction in call centre workers. *European Journal of Work and Organizational Psychology* 12, 366–392 (2003)
12. Van Dijk, P.A., Kirk-Brown, A.: Emotional labour and negative job outcomes: An evaluation of the mediating role of emotional dissonance. *Journal of Management and Organization* 12, 101–115 (2006)
13. Grandey, A.A.: When ‘The Show Must Go On’: Surface Acting and Deep Acting as Determinants of Emotional Exhaustion and Peer-Related Service Delivery. *Academy of Management Journal* 46, 86–96 (2003)
14. Parker, D.F., DeCotiis, T.A.: Organizational determinants of job stress. *Organizational Behavior and Human Performance* 32, 160–177 (1983)
15. March, J.G.: Exploration and Exploitation in Organizational Learning. *Organization Science* 2, 71–87 (1991)
16. Amabile, T.M.: The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology* 45, 357–376 (1983)
17. Sternberg, R.J.: *The Nature of Creativity: Contemporary Psychological Perspectives*. Cambridge University Press, Cambridge (1988)
18. Guilford, J.P.: Creativity. *Am. Psychiatr.* 5, 444–454 (1950)
19. Amabile, T.M.: *Creativity in context*. Westview, Boulder (1996)
20. Csikszentmihályi, M.: The Domain of Creativity. In: Runco, M.A., Albert, R.S. (eds.) *Theories of Creativity*, pp. 190–212. Sage, Newbury Park (1990)
21. Eisenberger, R., Aselage, J.: Incremental effects of reward on experienced performance pressure: positive outcomes for intrinsic interest and creativity. *Journal of Organizational Behavior* 30, 95–117 (2009)
22. Audia, P.G., Goncalo, J.A.: Past Success and Creativity over Time: A Study of Inventors in the Hard Disk Drive Industry. *Management Science* 53, 1–15 (2007)
23. Benner, M.J., Tushman, M.L.: Process management and technological innovation: A longitudinal study of the photography and paint industries. *Administrative Science Quarterly* 47, 676–706 (2002)
24. Benner, M.J., Tushman, M.L.: Exploitation, exploration, and process management: The productivity dilemma revisited. *Academy of Management Review* 28, 238–256 (2003)
25. Zhou, J., George, J.M.: When Job Dissatisfaction Leads To Creativity: Encouraging The Expression Of Voice. *Academy of Management Journal* 44, 682–696 (2001)
26. Prieto, I.M., Revilla, E., Rodriguez-Prado, B.: Managing the knowledge paradox in product development. *Journal of Knowledge Management* 13, 157–170 (2009)
27. Sosik, J.J., Godshalk, V.M.: Leadership styles, mentoring functions received, and job-related stress: a conceptual model and preliminary study. *Journal of Organizational Behavior* 21, 365–390 (2000)
28. Chin, W.W.: The Partial Least Squares Approach to Structural Equation Modeling. In: Marcoulides, G.A. (ed.) *Modern Methods for Business Research*, pp. 295–336. Lawrence Erlbaum Associates, Mahwah (1998)

The Impact of Creative Self-efficacy, IT Support, and Knowledge on Individual Creativity through Absorptive Capacity

Young Wook Seo¹, Seong Wook Chae², and Kun Chang Lee^{3,*}

¹ Software Engineering Center at NIPA, Seoul 138-711, Republic of Korea

² National Information Society Agency, Republic of Korea

³ SKK Business School and Department of Interaction Science,

Sungkyunkwan University, Seoul 110-745, Republic of Korea

{seoyy123, seongwookchae, kunchanglee}@gmail.com

Abstract. This study proposed a new individual creativity model that was divided into five main constructs: creative self-efficacy, IT support, individual knowledge, individual absorptive capacity, and individual creativity. We assumed that creative self-efficacy, IT support, and individual knowledge positively affect individual creativity through the mediating effect of individual absorptive capacity. Additionally, we examined the moderating effects of subjective well-being by dividing the sample into a high subjective well-being group and a low subjective well-being group. After collecting 606 valid questionnaires from IT companies in South Korea, we applied a structural equation modeling technique to analyze the data. Empirical results reveal the following: (1) creative self-efficacy, IT support, and individual knowledge influence individual creativity through individual absorptive capacity; and (2) subjective well-being moderates the relationship between the two constructs of the research model but does not moderate the influence of creative self-efficacy on individual absorptive capacity.

Keywords: Individual Creativity, Individual Absorptive Capacity, Creative Self-Efficacy, IT Support, Individual Knowledge, Subjective Well-Being.

1 Introduction

The patent war between Apple and Samsung Electronics over the design and operation of their smartphones and tablet computers is escalating and could turn into an all-out war. In today's fast-paced, knowledge-intensive, and Internet communication environment, information technology (IT) companies are placing more emphasis on creativity and innovation than ever before. Creativity has been defined as the production of novel and useful ideas in any domain, and innovation has been defined as the successful implementation of creative ideas within an organization [1]. IT companies are now under intense competition and are greatly concerned about the well-being of employees who need to manifest their creativity. However, relatively few studies have explored "individual creativity through absorptive capacity according to

* Corresponding author.

subjective well-being.” In the present article, we explore the effects of creative self-efficacy, individual knowledge, IT support, and individual absorptive capacity on individual creativity according to subjective well-being.

This research focuses on perceptions of individual creativity by asking the following three questions:

- (1) Do creative self-efficacy, individual knowledge, and IT support significantly influence individual absorptive capacity?
- (2) How do creative self-efficacy, individual knowledge, and individual absorptive capacity contribute to building individual creativity?
- (3) How does the level of subjective well-being moderate the relationship between individual creativity and its antecedents, such as creative self-efficacy, individual knowledge, IT support, and individual absorptive capacity?

2 Theoretical Background

2.1 Individual Creativity

Existing literature shows that the concept of creativity has expanded into diversified fields, including the arts, science, and business disciplines (e.g., [21, 22, 23]). Researchers have long recognized that creativity can refer to a person, process, product, or environmental response within even a single context [18]. Amabile’s [2, 3] componential theory of individual creativity includes three major components of individual (or small team) creativity, each of which is necessary for creativity in any given domain: expertise, creative-thinking skill, and intrinsic task motivation. Componential theory suggests that creativity is most likely to occur when people’s skills overlap with their strongest intrinsic interests — their deepest passions — and that creativity will be higher as the level of each of the three components increases [3]. Individual creativity at work is typically enacted in the context of a work team or group, where individual creative performance may be seen as a contribution to the team’s creative performance and achievement of team goals [15, 20].

2.2 Individual Absorptive Capacity

Absorptive capacity is receiving increasing attention as the emphasis on intangible assets such as knowledge, technology, and information strengthens, and prompt responses and innovation due to changing business environment become more important. The term “absorptive capacity” was first mentioned by Cohen and Levinthal [9], who defined it as an individual ability to acknowledge the value of new information and understand and apply it for a commercial purpose. Cohen and Levinthal argued that absorptive capacity motivates greater innovation and productivity [9]. The resource-based view (RBV), knowledge-based view (KBV), and dynamic capability theory provide a solid theoretical foundation for the term “absorptive capacity.”

Malhotra et al. [16] discovered that the intermediary role of absorptive capacity plays a positive role in improving operational efficiency and creating knowledge within a supply chain. Boynton et al. [5] described the importance of absorptive capacity and information technology, noting that absorptive capacity plays an intermediary role in utilizing information technology. Moreover, they argued that the absorptive capacity of a company provides a theoretical foundation for the profound understanding of information system utilization [5]. Zahra and George [25] explained that absorptive capacity is a necessary factor for the successful installation of a new information system.

3 Research Hypotheses

We developed a creativity model (Fig. 1) consisting of five main constructs: creative self-efficacy, IT support, individual knowledge, individual absorptive capacity, and individual creativity. The model assumes that creative self-efficacy, IT support, and individual knowledge have positive effects on individual creativity through the mediating effect of individual absorptive capacity. We also examined the moderating effects of subjective well-being by dividing the sample into a high group and a low group, hypothesizing that the influences of creative self-efficacy, IT support, and individual knowledge on creativity may differ between these two groups.

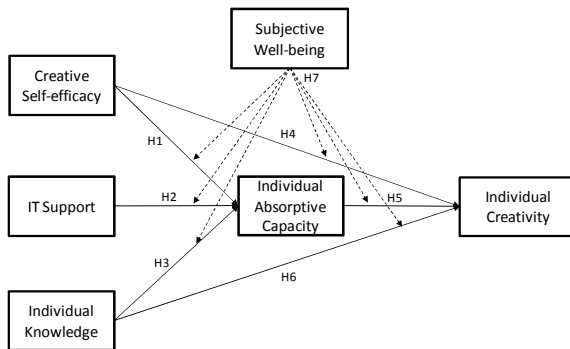


Fig. 1. Research Model

Working from Bandura's general definition of self-efficacy as targeted perceived capacity, Tierney and Farmer [24] defined creative self-efficacy as the belief that one has the ability to produce creative outcomes [4, 24]. Creative self-efficacy appears to provide momentum for creativity because strong efficacy beliefs enhance the persistence and coping efforts of individuals when encountering challenging situations [4, 24]. The critical role for IT lies in its ability to support communication, collaboration, and those searching for knowledge, in addition to its ability to enable collaborative learning [14]. IT used to support knowledge management may well revolutionize IT companies because effective IT support for knowledge management can serve as a competitive advantage and valuable professional aid to the absorptive

capacity of employees [14]. Zahra and George [25] claimed that exposure to external knowledge and experience plays a critical role in improving absorptive capacity. Therefore, the management environment allows individual knowledge to influence the absorptive capacity of organizational members. On the basis of these studies, we propose the following hypotheses.

H1: Creative self-efficacy positively contributes to individual absorptive capacity.

H2: IT support positively contributes to individual absorptive capacity.

H3: Individual knowledge positively contributes to individual absorptive capacity.

Creative self-efficacy has been used to explain creative performance above and beyond job self-efficacy [24]. In addition, individual differences in absorptive capacity are expected to affect learning, innovation, and productivity [10]. Deng et al. [10] conducted research on the relation between innovation and productivity in the field of IT engineering.

Amabile [2] identified both “domain-relevant skills” and “creativity-relevant skills” as being important for creativity. Domain-relevant skills include knowledge, expertise, technical skills, intelligence, and talent in the particular domain in which the problem-solver is working. As recognized by Mumford and Gustafson [17], knowledge plays a key role in creative achievement. According to Ford’s [13] theory of creative individual action, sense making, motivation, and knowledge and ability are three characteristics that influence individual creativity. On the basis of these studies, we propose the following hypotheses.

H4: Creative self-efficacy positively contributes to individual creativity.

H5: Individual absorptive capacity positively contributes to individual creativity.

H6: Individual knowledge positively contributes to individual creativity.

Well-being is generally classified into subjective well-being and psychological well-being. Subjective well-being is “a broad category of phenomena that includes people’s emotional responses, domain satisfactions, and global judgments of life satisfaction” [11]. It is related to the awareness of (dis)pleasure, satisfaction, and happiness [6] and can be measured by positive affect, negative affect, and satisfaction in life. Psychological well-being is an individual’s overall assessment on his/her life quality and is a multi-dimensional concept consisting of six dimensions: self-acceptance, personal growth, purpose in life, environmental mastery, autonomy, and positive relations with others [19].

Fierce competition among IT companies requires teams or individuals to generate creative and differentiated outputs in order to outperform competitors. Team members must have individual professional skills and knowledge for well-being in that environment. However, members of a team often have different individual levels of subjective well-being. Therefore, we divided the sample into a high group and a low

group according to their subjective well-being, and assumed that the influences of creative self-efficacy, IT support, and individual knowledge on creativity may differ between the two groups. We assumed that the effects of creative self-efficacy and IT support might be greater in the high-level subjective well-being group, whereas the effects of individual knowledge might be greater in the low-level group. Therefore, we proposed the following hypotheses.

H7: Differences in levels of subjective well-being influence the relationships between constructs.

H7-1: Creative self-efficacy has a greater relative effect on individual absorptive capacity in the high-level subjective well-being group than in the low-level group.

H7-2: IT support has a greater relative effect on individual absorptive capacity in the high-level subjective well-being group than in the low-level group.

H7-3: Individual knowledge has a greater relative effect on individual absorptive capacity in the low-level subjective well-being group than in the high-level group.

H7-4: Creative self-efficacy has a greater relative effect on individual creativity in the high-level subjective well-being group than in the low-level group.

H7-5: Individual absorptive capacity has a greater relative effect on individual creativity in the high-level subjective well-being group than in the low-level group.

H7-6: Individual knowledge has a greater relative effect on individual creativity in the low-level subjective well-being group than in the high-level group.

4 Experiments

4.1 Measurement and Survey

All items used to measure the constructs were adapted from reliable literature and measured on a 7-point Likert-type scale. We surveyed members of IT companies in South Korea faced with fierce competition to ensure that we included employees regularly engaging in rigorous creative activity (i.e., IT consultant, IT analyst, R&D researcher, system analyst, and so on). We received 661 responses from the research firm. If respondents supplied inconsistent or incomplete information or did not participate in creative activities in their position, their responses were eliminated from the dataset. This process yielded 606 useable questionnaires (91.68% response rate). We divided our sample into two groups based on the 7-point median value ($=4.00$) for subjective well-being (happiness) after averaging the six items that measured subjective well-being (happiness). The high-level subjective well-being (happiness) group (>4.00) included 484 respondents, the mean level of subjective well-being was 5.213, and the standard deviation was 0.890. The low-level group (≤ 4.00) included 122 respondents, the mean was 3.602, and the standard deviation was 0.476.

4.2 Results

We used SmartPLS 2.0 software to analyze our measurements and test our hypotheses. PLS is a structural-equation modeling tool that uses a component-based approach for estimation and places minimal restrictions on the sample size and residual distribution [8]. The individual item reliabilities, composite reliabilities (0.893-0.962), Cronbach’s alphas (0.819-0.951), and average variances extracted (0.682-0.835) for the constructs for each team indicated that all had acceptable levels of convergent validity and reliability. For discriminant validity, the AVE from the construct should be greater than the variance shared between the construct and other constructs in the model. In all cases, the AVE for each construct was larger than the correlation of that construct with all other constructs in the model (see Table 1).

Table 1. Convergent Validity and Discriminant Validity

Group	Construct	AVE	Composite Reliability	Cronbach’s Alpha	CS	ITS	IK	IAC	IC
All participants (N=606)	CS	0.795	0.921	0.871	0.892				
	ITS	0.835	0.962	0.951	0.385	0.914			
	IK	0.819	0.932	0.890	0.658	0.376	0.905		
	IAC	0.718	0.910	0.868	0.676	0.433	0.680	0.847	
	IC	0.809	0.944	0.921	0.658	0.355	0.619	0.654	0.899
High-level subjective well-being group (n=484)	CS	0.790	0.918	0.867	0.889				
	ITS	0.822	0.959	0.946	0.354	0.907			
	IK	0.795	0.921	0.871	0.638	0.342	0.892		
	IAC	0.700	0.903	0.857	0.649	0.420	0.648	0.837	
Low-level subjective well-being group (n=122)	CS	0.735	0.893	0.819	0.857				
	ITS	0.785	0.948	0.939	0.146	0.886			
	IK	0.823	0.933	0.892	0.541	0.143	0.907		
	IAC	0.682	0.896	0.845	0.605	0.188	0.626	0.826	
	IC	0.833	0.952	0.933	0.588	0.196	0.563	0.590	0.913

Note 1: TV= Task Variety, TA= Task Analyzability, IK= Individual Knowledge, ET= Exploitation, ER= Exploration, IC= Individual Creativity.

Note 2: AVE = Average variance extracted.

Note 3: The number on the diagonal denotes the square root of average variance extracted.

Note 4: For adequate discriminant validity, the diagonal elements should be greater than the corresponding off-diagonal elements.

As shown in Fig. 2, the structural model was tested using path coefficients and R² values. The path coefficients indicate the strength of the relationships between the dependent and independent variables, and the R² values represent the amount of variance explained by the independent variables.

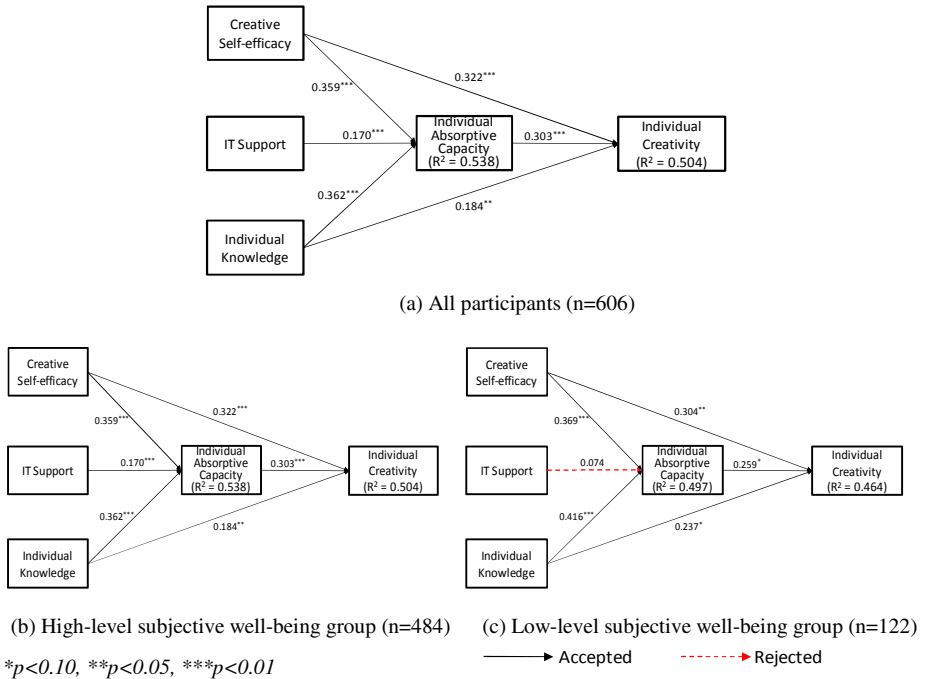


Fig. 2. Path Estimate by PLS Analysis

In this model, the R² values of all constructs exceeded the value (10%) recommended by Falk and Miller [12]. These results provide strong support for the relationships posited among the constructs.

When using the total sample of all participants, all hypotheses (H1-H6) were accepted. The set of hypotheses ranging from H7-1 to H7-6 was designed to investigate the moderating effects of subjective well-being. There were differences between the two groups (high and low) in testing H7-2. To examine the other hypotheses, we compared the path coefficients of the high-level subjective well-being group with those of the low-level group using the following formula suggested by Chin et al. [7]:

$$t_{ij} = \frac{p_1 - p_2}{\sqrt{\frac{(n_1 - 1) \times SE_1^2 + (n_2 - 1) \times SE_2^2}{n_1 + n_2 - 2} \times \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

P_i: ith path coefficient, N_i: ith sample size, SE_i: ith standard error, degree of freedom for t_{ij}: n₁ + n₂ - 2

The analysis revealed that subjective well-being moderated the relationship between the two constructs, H7-1 (t_{spooled} = -1.591), H7-2 (t_{spooled} = 15.500), H7-3 (t_{spooled} = -9.292), H7-4 (t_{spooled} = 2.686), H7-5 (t_{spooled} = 5.521) and H7-6 (t_{spooled} = -5.825). Thus, all hypotheses (H7-2 to H7-6) on moderating effects were supported with the exception of H7-1.

5 Discussion and Conclusion

We implemented an empirical test of responses to questionnaires collected from members of IT companies in order to analyze the impact of creative self-efficacy, IT support, and individual knowledge on individual creativity through individual absorptive capacity. We also examined differences in the effects of the independent variables (creative self-efficacy, IT support, and individual knowledge) on subjective well-being between the two groups.

Employees in IT companies are supported by the information systems of their companies and manifest their individual capacity by sharing and utilizing their knowledge. These employees perform a greater number of creative tasks than employees of other industries. This empirical study has several key findings regarding the effect of subjective well-being. First, IT companies should place more emphasis on the well-being (happiness) of their employees so as to create an environment in which they can manifest a higher level of individual creativity. Such efforts will bring greater levels of innovation and benefits to the companies. Second, in the low-level subjective well-being group, the effect of individual knowledge on individual absorptive capacity and creativity were strong, but IT support did not affect individual absorptive capacity. Hence, IT companies should create an environment that elevates the well-being of the employees but, at the same time, should set a management strategy to help their IT support positively affect individual absorptive capacity. In other words, IT companies should utilize their IT-related information system to help individuals manifest their capacity in carrying out creative tasks by accumulating, sharing, and utilizing the professional knowledge of individuals and the organization. Third, IT companies may be able to elevate individual creativity only when they create an environment that elevates the individual absorptive capacity. These companies should increase the level of individual creative self-efficacy and individual knowledge and form an environment in which IT support and its utilization can be performed well by the organization. Additionally, companies need to implement educational training programs that can enhance individual creativity. IT companies should help employees manifest their creativity in occupational tasks by elevating their quality of life.

This paper has limitations that should be addressed in future research. First, we did not consider antecedents of individual creativity that may be important. Future research should probe how knowledge sharing, exploitation, and exploration influence individual creativity. Second, we did not consider many measurement items of well-being. Future researchers should explore how subjective well-being (such as quality of life or life satisfaction) and psychological well-being (such as self-acceptance, personal growth, purpose in life, environmental mastery, autonomy, or positive relations with others) may moderate the relationships among the constructs in the model.

Acknowledgments. This study was supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0).

References

1. Amabile, T.M., Conti, R., Coon, H., Lazenby, J., Herron, M.: Assessing the work environment for creativity. *Academy of Management Journal* 39, 1154–1184 (1996)
2. Amabile, T.M.: A Model of Creativity and Innovation in Organizations. In: Staw, B.M., Cummings, L.L. (eds.) *Research in Organizational Behavior*, vol. 10, pp. 123–167. JAI Press, Greenwich (1988)
3. Amabile, T.M.: Motivating Creativity in Organizations. *California Management Review* 40(1), 22–26 (1997)
4. Bandura, A.: *Self-efficacy: The exercise of control*. Freeman, New York (1997)
5. Boynton, A.C., Zmud, R.W., Jacobs, G.C.: The influence of IT management practice on IT use in large organizations. *MIS Quarterly* 18(3), 299–318 (1994)
6. Burns, R.A., Machin, M.A.: Investigating the Structural Validity of Ryff's Psychological Well-Being Scales across Two Samples. *Social Indicators Research* 93(2), 359–375 (2009)
7. Chin, W.W., Marcolin, B.L., Newsted, P.R.: A Partial Least Squares Latent Variable Modeling Approach for measuring Interaction Effect: Results from a Monte Carlo Simulation Study and Voice Mail Emotion/adoption Study. In: 17th International Conference on Information Systems, OH, pp. 21–41 (1996)
8. Chin, W.W.: The partial least squares approach to structural equation modeling. In: Marcoulides, G.A. (ed.) *Modern Methods for Business Research*. Lawrence Erlbaum, Mahway (1988)
9. Cohen, W.M., Levinthal, D.A.: Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35(1), 128–153 (1990)
10. Deng, X., Doll, W.J., Cao, M.: Exploring the absorptive capacity to innovation/productivity link for individual engineers engaged in IT enabled work. *Information & Management* 45, 75–87 (2008)
11. Diener, E., Suh, E.M., Lucas, R.E., Smith, H.L.: Subjective Well-Being: Three Decades of Progress. *Psychological Bulletin* 125(2), 276–302 (1999)
12. Falk, R.F., Miller, N.B.: *A premier for soft modeling*. The University of Akron, Akron (1988)
13. Ford, C.M.: A Theory of Individual Creative Action in Multiple Social Domains. *Academy of Management Review* 21(4), 1112–1142 (1996)
14. Gottschalk, P.: Predictors of Information Technology Support for Inter-organizational Knowledge Management: Lessons Learned from Law Firms in Norway. *Knowledge and Process Management* 8(3), 186–194 (2001)
15. Hirst, G., van Dick, R., van Knippenberg, D.: A social identity perspective on leadership and employee creativity. *Journal of Organizational Behavior* 30, 963–982 (2009)
16. Malhotra, A., Gosain, S., El Sawy, O.A.: Absorptive Capacity Configurations in Supply Chains: Gearing for Partner-Enabled Market Knowledge Creation. *MIS Quarterly* 29(1), 145–187 (2005)
17. Mumford, M.D., Gustafson, S.B.: Creativity Syndrome: Integration, Application and Innovation. *Psychological Bulletin* 103, 27–43 (1988)
18. Rhodes, M.: An analysis of creativity. *Phi Delta Kappan* 42, 305–310 (1961)
19. Ryff, C.D., Keyes, C.: The structure of psychological well-being revisited. *Journal of Personality and Social Psychology* 69, 719–727 (1995)
20. Shin, S., Zhou, J.: When is educational specialization heterogeneity related to creativity in research and development teams? Transformational leadership as a moderator. *Journal of Applied Psychology* 92, 1709–1721 (2007)

21. Stumpf, H.: Scientific creativity: A short overview. *Educational Psychology Review* 7, 225–241 (1995)
22. Tang, P.C., Leonard, A.R.: Creativity in art and science. *Journal of Aesthetic Education* 19, 5–19 (1985)
23. Taylor, C.W.: Approaches to and definitions of creativity. In: Sternberg, R.J. (ed.) *The Nature of Creativity: Contemporary Psychological Perspectives*, pp. 99–121. Cambridge University Press, Cambridge (1988)
24. Tierney, P., Farmer, S.M.: Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal* 45(6), 1137–1148 (2002)
25. Zahra, S.A., George, G.: The Net-Enabled Business Innovation Cycle and the Evolution of Dynamic Capabilities. *Information Systems Research* 13(2), 147–150 (2002)

An Empirical Analysis of Leadership Styles and Their Impact on Creativity: Emphasis on Korean ICT Companies

Nam Yong Jo¹, Dae Sung Lee¹, and Kun Chang Lee^{2,*}

¹ SKK Business School Sungkyunkwan University, Seoul 110-745, Republic of Korea
² Professor at SKK Business School WCU Professor at Department of Interaction Science
Sungkyunkwan University Seoul 110-745, Republic of Korea
{namyong.jo, leeds1122, kunchanglee}@gmail.com

Abstract. This research builds and tests a theoretical model linking empowering leadership and servant leadership with several intervening variables. Using survey data from professional employees working for Korean ICT companies, we compared both leadership styles' influences on several mediating variables and found that 1) both leadership styles negatively affect employees' task stress but positively affect trust in both leaders and other employees 2) they also inversely affect intrinsic motivation, in that empowering leadership positively affects motivation but servant leadership negatively affects it 3) both types of trust have a positive influence on employees' creativity 4) intrinsic motivation has a strong positive effect on employees' creativity but 5) no significant effect of stress on creativity was found.

Keywords: Empowering Leadership, Servant Leadership, Task Stress, Trust, Intrinsic Motivation, Creativity.

1 Introduction

Competition in the ICT industry is becoming severe, especially in the software domain. For example, two of the world's major Smartphone manufacturers, Apple and Samsung Electronics, are now battling each other in lengthy, worldwide patent lawsuits. Apple insists that the Galaxy Tab copied the designs of its products, while Samsung insists iPhone 3rd and 4th generations and iPad 2 devices infringed upon patents relating to wireless networking technology. Their patents, such as UX designs of devices and wireless technologies are creative software assets produced through teams or employees' creative works. The ICT industry, particularly the software domain, which is characterized by fast-growth and strong competition, is one of the most interesting fields in which to study creativity; however, the subject has not received much attention from researchers.

Organizational researchers believe that the relationship between leaders and followers affects followers' creativity significantly. For instance, Tierney et al. (1999),

* Corresponding author.

and Scott & Bruce (1994) found that effective leader-member exchange (LMS) relationships are positively associated with employees' creativity. Studies also have provided evidence for a positive relationship between controlling leadership and employee creativity (Amabile et. al., 2004; Madjar et al., 2002; Oldham & Cummings, 1996; Tierney & Farmer, 2002: 2004). Together with leadership styles, mediating variables affecting employees' creativity also have been found. We reviewed these studies and then proposed a model to compare, concurrently, two types of leadership followed by mediating variables: trust in the leader, and trust in other members, task stress and intrinsic motivation affecting employees' creativity. From our model, the influence of task stress and trust on an individual's creativity remain inconclusive. However, stress and trust are, for all practical purposes, psychological factors, as employees in the ICT industry tend to experience more stress in the highly competitive domain of software development, with its intangible and creative characteristics.

Thus, for the empirical purpose of this study, this paper is organized as follows. Section 2 contains a brief overview of studies related to our model and presents our hypotheses. In Section 3, the methods of this study and statistical analyses are set forth, and the results and concluding remarks are presented in the final sections.

2 Theory and Hypotheses

There has been discussion in the creativity literature about different styles of leadership and their influence on employees' creativity (Amabile, 2004; George & Zhou, 2007; Shalley & Gilson, 2004). Some creativity researchers believe that leadership styles fundamentally address the nature of creative works. Typically, transformational leadership and empowering leadership are understood to influence individuals' creativity by actively encouraging employees to use self-direction and self-motivation (Pearce & Sims, 2002). LMX theory is also used to understand the relationship between servant-leader and follower, but servant leadership, which is overlapping somewhat with empowering leadership, has not received much attention so far. Along with these studies, which have provided evidence for a positive relationship between supportive leadership and creativity, we are going to discuss empowering leadership and servant leadership in this study.

There are some characteristics of servant and empowering leadership that are shared, such as empowering and developing employees by delegating authority to increase intrinsic motivation, accentuating accountability by giving people clear goals, and requiring managers to share knowledge and information to ensure that employees develop the necessary skills. However, certain characteristics of servant leadership, such as expressing humility and authenticity, emphasizing interpersonal acceptance, and demonstrating stewardship, are not explicitly formulated as part of empowering leadership. Thus, in this study we propose to compare the effect of these two leadership styles on individuals' creativity.

2.1 Leadership and Task Stress

Task stress is conceptualized as an individual's feeling of uncertainty and of being threatened when work demands exceed the individual's capacity to perform optimally

(Edwards, 1992). In such situations, individuals feel as if they have lost control of their work environment and that they can never escape it. Worse, employees bring the work and their worries about it home. To decrease feelings of stress, managers can support employees to increase certainty and limit perceived threats in the working environment. Researchers have found that employee's high-quality LMX relationships with their leaders can serve this purpose (e.g., Beehr, King, & King, 1990; Cohen & Wills, 1985; Ganster, Fusilier, & Mayes, 1986; Viswesvaran, Sanchez, & Fisher, 1999). However, few investigators have directly examined the leadership and stress relationship. Thus, we propose:

Hypothesis 1a: Servant leadership has a negative influence on task stress

Hypothesis 1b: Empowering leadership has a negative influence on task stress

2.2 Leadership and Trust

Trust is defined as a willingness to depend on another party (Mayer et al., 1995) and the expectation that the other party will reciprocate if one cooperates. Employee's perception of their leader's ability or competence is an essential factor in the level of trust within an organization, because employees are unlikely to trust in their leader unless the leader is capable of fulfilling the leadership role (Whitener et al., 1998). Trust in the leader leads to such positive outcomes as organizational citizenship behaviors, satisfaction, and performance (e.g. Jung and Avolio, 2000; Pillai et al., 1999). Trust can also exist at both the organizational and inter-organizational levels and it is likely to develop differently in relation to team members, team leaders, and to the organization. If interaction with team members occurs more frequently, these are likely to be fundamentally different compared to the relationship with team leaders. Thus we propose the alternative hypotheses:

Hypothesis 2a: Servant leadership has a positive influence on trust in the leader

Hypothesis 2b: Empowering leadership has a positive influence on trust in the leader

Hypothesis 3a: Servant leadership has a positive influence on trust in the members

Hypothesis 3b: Empowering leadership has positive influence on trust in the members

2.3 Leadership and Intrinsic Motivation

Empowerment is a motivational concept that aims to facilitate a proactive, self-confident attitude among employees and to give them a sense of personal power (Conger, 2000). Empowering leadership behavior includes such behaviors as: encouraging self-directed decision making, information sharing and coaching for innovative performance (Konczak, Stelly, & Trusty, 2000). The servant-leader believes in the intrinsic value of each individual; it is all about recognition, acknowledgement, and the realization both of an individual's current capabilities and what s/he can still learn (Greenleaf, 1998). Thus, we propose the following hypotheses.

Hypothesis 4a: Servant leadership has a positive influence on intrinsic motivation

Hypothesis 4b: Empowering leadership has a positive influence on intrinsic motivation

2.4 Stress and Creativity

According to distraction arousal theory (Byron et al. 2010), stress decreases creative performance. If people devote their limited resources to stressors, there remain fewer cognitive resources available for other productive tasks. The reduction of cognitive resources results in using a narrow intentional focus (Eysenck, 1995). That simple cognitive strategy is likely to result in the production of more ordinal, less creative ideas (Baron, 1986; Drwal, 1973). In short, stress decreases creativity by demanding cognitive resources that are then not available for creative thinking and instead result in the use of simpler cognitive strategies that are likely to undermine creativity.

However, there are other researchers who insist that stress increases creativity. According to these studies, stressors increase arousal, so that people use creative thought and are motivated to derive solutions (Anderson et al., 2004). In other words, when people are exposed to stressors, they may engage in focused problem-solving strategies, thereby enhancing creativity (Bunce & West, 1994), which then provides the cognitive stimulation and motivating arousal that are needed for creative thinking (Andrews & Farris, 1972; Pelz, 1988). Considering these opposing studies, but advocating the former study, we propose the following hypothesis:

Hypothesis H5: Task stress has a negative influence on employees' creativity

2.5 Trust and Creativity

Despite the general consensus that mutual trust is a key success factor for innovative developments, prior research on the influence of trust on creativity remains largely inconclusive. For instance, some researchers (Jehn, 1995; Simons & Peterson, 2000; Dakhli & De Clercq, 2004) examined the role of mutual trust in increasing team creativity, but some recent studies (Chen, Chang & Hung, 2008) do not find that there is a positive influence of mutual trust on team creativity. However, it is likely that trust arises in more open, supportive, tolerant environments (Carnevale & Probst, 1998). Still, previous research does not fully support the existence of a definite relationship between trust and creativity. Thus, our research aimed to further investigate the role of trust on creativity.

Hypothesis H6: Trust in the leader has a positive influence on employees' creativity

Hypothesis H7: Trust in members has a positive influence on employees' creativity

2.6 Intrinsic Motivation and Creativity

Intrinsic motivation refers to the desire to expend effort based on interest in and enjoyment of the work being performed (Amabile, 1996). Zhang and Bartol (2010) broadly

cite Amabile's (1983) componential conceptualization of creativity, then insist that intrinsic motivation is one of the most important and powerful influences on employee creativity. Grant and Berry (2011) identified three interrelated psychological mechanisms which stimulate creativity when employees are intrinsically motivated: first, they are positively affected (Silvia, 2008); second, their curiosity and interest in learning will enhance their cognitive flexibility, willingness to take risks, and openness to complexity, which in turn will expand their access to ideas and solutions (Ganne & Deci, 2005); third, they are promoted creativity by encouraging persistence. Thus, intrinsic motivation enhances engagement and energy for sustaining effort, thereby increasing the amount of time that they are willing to work on their task (Fredrickson, 1998). Thus, our final hypothesis addresses the effect of intrinsic motivation on creativity

Hypothesis H8: Intrinsic motivation has a positive influence on employees' creativity

3 Methods

3.1 Questionnaire Survey and Sample Statistics

The samples were collected by a research agency in Korea. The respondents work for more than three hundred ITC companies. The numbers of respondents are 365 but we purposely included only 249 respondents (n. male=191, female=58), because we excluded the respondents who are leaders.

3.2 Reliability and Confirmatory Factor Analyses

Preliminary analysis

Seven constructs of creativity: trust in the leader; trust in the member; task stress; empowering leadership; servant leadership, and intrinsic motivation were used in this analysis. The survey items of each construct were reliable, with Cronbach's alpha values greater than 0.7. The validities of the survey items were tested using a principal component analysis with varimax rotation. The first factor, creativity, explained 32.7% of the total variance; the second factor, trust in the leader, accounted for 13.31%. The total variance explained by the seven factors was 72.1%, as shown in Table 1. From these results, we concluded that the survey items were statistically valid.

The measurement model

To confirm the reliability and validity of the measurement data, we performed a confirmatory factor analysis. In the reliability test, all of the composite reliabilities except stress were greater than 0.7, and the AVEs (Average Variance Extracted) were greater than 0.5 [35]. For the validity test, the correlation between two factors was less than the square root of the AVE value of each factor [35]. Therefore, as shown in Table 2, these data were reliable and valid except for the stress variable.

Table 1. Results of reliability and factor analyses

Survey item	Cronbach's α	CRT	LTR	STR	ELS	MTV	MTR	SLS
creativity3	0.940	0.848						
creativity4		0.825						
creativity5		0.820						
creativity7		0.793						
creativity6		0.748						
creativity2		0.741						
creativity1		0.695						
creativity8		0.657						
creativity9		0.642						
leader_trust2	0.942		0.834					
leader_trust4			0.826					
leader_trust5			0.823					
leader_trust8			0.794					
leader_trust1			0.793					
leader_trust6			0.777					
leader_trust3			0.770					
leader_trust7		0.769						
Stress7	0.933			0.887				
Stress8				0.872				
Stress6				0.869				
Stress3				0.845				
Stress2				0.819				
Stress1				0.782				
Stress5				0.760				
Stress4			0.758					
empower_leadership3	0.912				0.731			
empower_leadership1					0.727			
empower_leadership2					0.718			
empower_leadership5					0.700			
empower_leadership6					0.684			
empower_leadership4					0.620			
motivation4	0.863				0.765			
motivation1					0.746			
motivation3					0.744			
motivation5					0.714			
motivation2					0.657			
member_trust3	0.933					0.840		
member_trust2						0.825		
member_trust4						0.822		
member_trust1						0.761		
servant_leadership3	0.892						0.787	
servant_leadership4							0.758	
servant_leadership2							0.708	
servant_leadership3							0.629	
Eigen value		14.389	5.856	4.572	1.998	1.946	1.687	1.257
Variance explained		32.702	13.310	10.390	4.541	4.423	3.834	2.857
Total variance explained (%)		32.702	46.012	56.402	60.943	65.367	69.201	72.058

* Note: CRT: Creativity, LTR: Trust on Leader, STR: Task Stress, ELS: Empowering Leadership, MTV: Motivation, MTR: Trust on members, SLS: Servant leadership.

The structural model

This study's hypotheses were tested using SmartPLS 2.0 with the bootstrapping procedure, and all of the hypotheses were accepted at a 99% confidence level. As is shown in Figure. 1, hypothesis H2a, H2b, H3a, H3b, H4b, H6, H7, and H8 were significantly accepted.

Table 2. Correlations of latent variables and AVEs

Construct	Composite Reliability	CRT	LTR	STR	ELS	MTV	MTR	SLS
Creativity	0.950	<u>(0.825)</u>						
Leader Trust	0.953	0.352	<u>(0.846)</u>					
Stress	0.516	-0.271	-0.130	<u>(0.415)</u>				
Empowering Leadership	0.932	0.532	0.586	-0.153	<u>(0.834)</u>			
Motivation	0.902	0.614	0.226	-0.223	0.441	<u>(0.793)</u>		
Member Trust	0.953	0.573	0.372	-0.153	0.488	0.413	<u>(0.913)</u>	
Servant Leadership	0.925	0.366	0.597	-0.128	0.640	0.120	0.392	<u>(0.869)</u>

Note: Values on the underlined diagonal are the square roots of the AVEs.

The measurement model

To confirm the reliability and validity of the measurement data, we performed a confirmatory factor analysis. In the reliability test, all of the composite reliabilities except stress were greater than 0.7, and the AVEs (Average Variance Extracted) were greater than 0.5 [35]. For the validity test, the correlation between two factors was less than the square root of the AVE value of each factor [35]. Therefore, as shown in Table 2, these data were reliable and valid except stress variable.

Table 3. Correlations of latent variables and AVEs

Construct	Composite Reliability	CRT	LTR	STR	ELS	MTV	MTR	SLS
Creativity	0.950	<u>(0.825)</u>						
Leader Trust	0.953	0.352	<u>(0.846)</u>					
Stress	0.516	-0.271	-0.130	<u>(0.415)</u>				
Empowering Leadership	0.932	0.532	0.586	-0.153	<u>(0.834)</u>			
Motivation	0.902	0.614	0.226	-0.223	0.441	<u>(0.793)</u>		
Member Trust	0.953	0.573	0.372	-0.153	0.488	0.413	<u>(0.913)</u>	
Servant Leadership	0.925	0.366	0.597	-0.128	0.640	0.120	0.392	<u>(0.869)</u>

Note: Values on the underlined diagonal are the square roots of the AVEs.

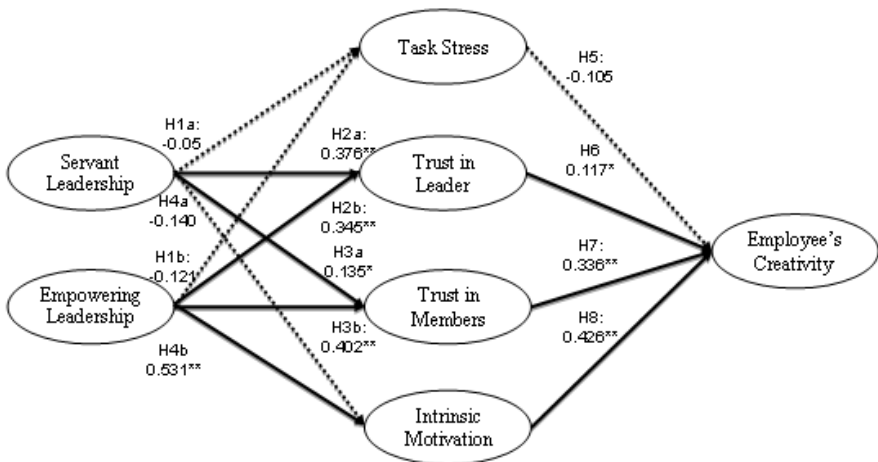


Fig. 1. Results of the structural model

The structural model

This study's hypotheses were tested using SmartPLS 2.0 with the bootstrapping procedure, and all of the hypotheses were accepted at a 99% confidence level. As is shown in Figure. 1 hypothesis H2a, H2b, H3a, H3b, H4b, H6, H7, and H8 were significantly accepted.

4 Discussion

This paper addressed the empowering leadership and servant leadership styles and psychological factors that affect employees' creativity while working in the ITC industry using the PLS. From the statistical analysis, we reached the following conclusions.

First, our leadership styles of interest, servant leadership and empowering leadership, did not correlate significantly with task stress (H1a, H1b). In other words, supportive leaderships do not relieve employees' immediate task stress. This result does not correspond to previous research which insisted that supportive leadership styles and good quality of LMS relieve followers' stress (e.g., Beehr, King, & King, 1990; Cohen & Wills, 1985; Ganster, Fusilier, & Mayes, 1986; Viswesvaran, Sanchez, & Fisher, 1999). The survey items for measuring stress consist of the individual's rating of anxiety and time pressure. Thus, further analysis needs to be conducted. Second, both leadership styles positively affect trust in the leader (H2a, H2b) as well as trust in other members in the organization (H3a, H4b). There is a notable difference between the two leadership styles' influences on trust-target. In other words, empowering leadership significantly influences trust in members but not trust in the leader. Thus, we may conclude that when employees have more self-directed decision making they tend to trust their colleagues. Third, empowering leadership has a significant positive influence on intrinsic motivation (H4b), while servant leadership does not (H4a). This result suggests to us that, although both leadership styles have many similar aspects, empowering leadership is more likely to inspire employees' intrinsic motivation. Fourth, task stress does not significantly affect employees' creativity (H5). According to previous research, task stress has a curvilinear relationship with creativity; moreover, another study asserts that task stress is affected by task complexity. Thus, more detailed analysis of this relationship is needed. Fifth, both trust in the leader and member have a positive influence on employees' creativity (H6, H7), and the latter influence is significantly stronger. We conclude from this result that individual creativity is more closely related to his/her co-worker than leader, due to the fact that there are more discussions and cooperation among colleagues than with the leader. Thus, this result is intuitively assumed. Finally, intrinsic motivation positively influences an employee's creativity, in keeping with most previous research and our last hypothesis (H8).

5 Concluding Remarks

We proposed a research model to discuss the relationship between leadership styles and employees' creativity. Since the ICT industry is a fast-growing and highly competitive environment, individuals in the industry are supposed to be subjected to greater stress; thus, if this stress is relieved, a higher quality of LMX is expected. While previous studies have tried to understand leadership styles, such as empowering leadership and transformational leadership, the influences of servant leadership, comparison among leadership styles, and mediating variables toward employee's creativity have not been well-addressed thus far. For these reasons, an empirical analysis was performed with survey data collected from 249 under-senior manager employees of Korean ICT companies.

Using PLS analysis with the survey data, we found that 1) neither servant leadership nor empowering leadership significantly affects task stress 2) both servant and empowering leadership have positive influences on both trust in the leader and trust in the member 3) empowering leadership positively affects intrinsic motivation but servant leadership does not 4) task stress does not significantly affect employees' creativity, and 5) both trust in the leader and members have positive influences on employees' creativity. Finally, intrinsic motivation has a positive influence on employees' creativity. From a detailed review of the results, we can conclude that more creative individuals trust in members who are encouraged through empowering leadership rather than servant leadership. Moreover, empowering leadership better motivates employees intrinsically.

In conclusion, we hope this study helps improve understanding, not only for researchers in the creativity field of study, but also organizational leaders in the ICT industry, about the influence among leadership styles, trust and stress on employees' creativity. This study has some limitations. For example, the influence of task stress is not clearly analyzed based on previous research. Further analysis to discover if there is a curvilinear relationship between task stress and employees' creativity is necessary. Task complexity as a control variable also needs to be examined. These factors will be analyzed in future work.

Acknowledgments. This research was supported by WCU(World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0).

References

1. Grant, M.G., Berry, J.W.: The Necessity of Others is the Mother of Invention: Intrinsic and Prosocial Motivations, Perspective Taking, and Creativity. *Academy of Management Journal* 54, 73–96 (2011)
2. Zhang, X., Bartol, K.M.: Linking Empowering Leadership and Employee Creativity: The Influence of Psychological Empowerment, Intrinsic Motivation, and Creative Process Engagement. *Academy of Management Journal* 53, 107–128 (2010)

3. Kenneth, J., Harris, K., Michele, K.: Too Much of a Good Thing: The Curvilinear Effect of Leader-Member Exchange on Stress. *Journal of Social Psychology* 146, 65–84 (2006)
4. Byron, K., Khazanchi, S., Nazarian, D.: The relationship between stressors and creativity: A meta-analysis examining competing theoretical models, vol. 95, pp. 201–212 (2010)
5. Bartram, T., Casimir, G.: A Bayesian method for the induction of probabilistic networks from data. *Leadership & Organization Development Journal* 28, 4–19 (2007)
6. Burke, C.S., Sims, D.E., Lazzara, E.H., Salas, E.: Trust in leadership: A multi-level review and integration. *The Leadership Quarterly* 18, 606–632 (2007)
7. Bidault, F., Castello, A.: Trust and Creativity: Understanding the Role of Trust in Creativity Oriented Joint Developments. *R&D Management* 39, 259–270 (2009)
8. Bidault, F., Castello, A.: Trust and Creativity: Identifying the Role of Trust in Creativity-Oriented Joint-Developments. *European School of Management and Technology, Working paper*, No. 08-010

A Physiological Approach to Creativity under Stress and Non-stress Conditions

Dae Sung Lee¹, Nam Yong Jo¹, and Kun Chang Lee²

¹ PhD Candidate, SKK Business School, Sungkyunkwan University,
Seoul 110-745, Republic of Korea

² Professor of MIS and WCU Professor of Creativity Science,
SKK Business School and Department of Interaction Science,
Sungkyunkwan University, Seoul 110-745,
Republic of Korea

{leeds1122, namyong.jo, kunchanglee}@gmail.com

Abstract. We examined creativity using a physiological approach, which requires a controlled experiment. We measured subjects' Galvanic Skin Response (GSR) and electrocardiogram (ECR) during a creativity task, and inserted two stress manipulations into a specific group. In addition, a questionnaire survey was given to subjects as a supplementary measurement. Physiological signals and the perceived stress scale (PSS) indicated whether or not the stress group had mental stress in the experiment. Our results showed self-reported creativity did not correspond to assessment by experts in that it might be influenced by our stress manipulation in the experiment. Finally, we did not find that there was a relation between creativity and stress for our experiment. For future studies of creativity under stress, we suggest that researchers should consider other factors that might influence stress and creativity with various approaches, such as physiological and bioinformatical approaches, in order to shed light on inconsistent findings.

Keywords: Stress, creativity, physiological signals.

1 Introduction

Psychological, organizational, and educational literatures have examined the relationship between stress and creativity with considerable attention. This study used as physiological approach to analyze the effect of stress on creativity. For this experiment, we adopted two stress manipulations (threat of shock and performance feedback) for our stress group. The physiological signals used in our experiment were Galvanic Skin Response (GSR) and electrocardiogram (ECG). A total of 31 subjects were employed in this study. The mean age was 22 years. Some of the subjects were randomly assigned to stress manipulations. Consequently, the subjects were divided into two groups (stress group, $n = 16$; non-stress group, $n = 15$). The subjects were instructed to start mediating for seven minutes in order to acquire baseline data from GSR and ECG signals. And then creativity tasks were delivered to the subjects. The

subjects were asked to generate creative ideas for three different toys in 10 minutes, during which data from GSR and ECG signals were obtained. After the physiological experiment, the subjects were requested to complete the questionnaire survey, which consisted of stress and creativity items.

For experiment analyses, we explored whether our stress manipulations coincide with self-reported stress (using the Perceived Stress Scale [PSS]). Moreover, we investigated whether creativity assessed by experts correspond to self-reported creativity. If there should be a discrepancy between the two rating methods, we would search the implicit meanings of that. First the relationship of stress and creativity was examined. We start off reviewing previous studies and finalize our research by showing our interpretations.

2 Theoretical Background

2.1 Stress and Creativity

In previous studies, the influence of stress on performance may have been represented by an inverted U-shaped (curvilinear) relationship. Namely, an appropriate amount of stress might be beneficial to help maintain concentration for a task, through excitement. If the amount should be beyond a threshold point stress might have negative effects on performance. For the purpose of our study, we define stress as distress, or stress that has reached a level at which it exhibits negative effects on cognitive performance, such as on a creative task. In our experiment, we use two stress manipulations (threat of shock and performance feedback) on a specific group. Studies employing a stimulus approach are concerned with stressors; that is, physical or psychological conditions that necessitate an adaptive response [13, 16]. Creativity is the production of ideas, solutions, or products that are novel and useful in a given situation [1]. We use a product-based approach to creativity, which is the most widely accepted approach to studying creativity. Therefore, our subjects are instructed to complete creativity tasks through creative thinking in order to investigate the relationship between stress and creativity.

2.2 Assessment of Stress

Psychological Assessment

Stress response can be measured and evaluated in terms of perceptual, behavioral, and physical responses. The evaluation of perceptual responses to a stressor involves subjective estimations and perceptions. Self-reported questionnaires are the most common instruments used to measure stress [8]. Representative measures are the Perceived Stress Scale (PSS) [7], the Life Events and Coping Inventory (LECI) [10], and the Stress Response Inventory (SRI) [14]. The physical response to stress has two components: a physiological response indicative of central-autonomic activity and a biochemical response involving changes in the endocrine and immune systems [8].

Physiological Assessment

Stress induces a change in autonomic functioning [26]. Stress influences blood pressure and heart rate, reflecting a predominance of sympathetic nervous system activity [21]. Heart rate variability (HRV) is beat-to-beat variation in heart rate, and it has recently been used as a biomarker of Autonomic Nervous System (ANS) activity associated with mental stress [27]. HRV analysis is generally divided into two categories: time-domain and frequency-domain method. Time-domain analysis of HRV involves quantifying the mean or standard deviation of RR intervals. Frequency-domain analysis involves calculating the power of the respiratory-dependent high frequency (HF) and low frequency (LF) components of HRV. In this study, we select the standard deviation of RR intervals (SDNN) and LF/HF ratio as ECG information. Mental stress is reported to evoke a decrease in the high-frequency component and an increase in the low-frequency component of HRV [3]. Therefore, LF/HF ratio increases if mental stress arises. On the other hand, a decrease of SDNN is related to mental stress. GSR is a measure of the electrical resistance of the skin. A transient increase in skin conductance is proportional to sweat secretion [9]. When an individual is under mental stress, sweat-gland activity is activated and increases skin conductance. Since the sweat glands are also controlled by the SNS (Sympathetic Nervous System), skin conductance acts as an indicator for sympathetic activation due to the stress reaction.

3 Method

3.1 Participants

Thirty seven healthy subjects recruited from undergraduate students at a Korean university participated in this experiment. Prior to the experiment, the subjects were given written and verbal information explaining the experimental procedures. We confirmed through interviews that none of the subjects used medication for hypertension or any other cardiovascular disease and they were all free of any nervous or other psychological disorder. We received written informed consent from all participants and each subject was paid 20,000 Korean won for this participation. Among them, six subjects with corrupted data were eliminated from this study. A total of 31 subjects (22 men and 9 women) were employed in this study. The mean age was 22 years (range of 18 - 26 years). Some of the subjects were randomly assigned to stress manipulations. Consequently, the subjects were divided into two groups (stress group, $n = 16$; non-stress group, $n = 15$).

3.2 Experimental Procedure

Before the measurements, the subjects were instructed to cleanse their hands and remove all accessories from their body. Then, the subjects were asked to sit comfortably and keep their left hand still when the experiment started. Each subject was asked to attach two GSR electrodes to the index and middle finger of the left hand and place three ECG electrodes on the chest and abdomen. This experiment used

a Biopac MP100 series for the measuring and an AcqKnowledge 4.1 for the analysis. After GSR and ECG signals had showed normal waves, the subjects were instructed to start meditating for seven minutes in order to acquire baseline data from GSR and ECG signals. And then creativity tasks were delivered to the subjects. The subjects were asked to generate creative ideas for three different toys in 10 minutes. At the same time, the subjects were shown examples used by Smith et al. [25]. In the course of task implementation, GSR and ECG signals were measured for both the stress and non-stress group. However, stress manipulations were inserted into the stress group. After the physiological experiment, the subjects were requested to complete the questionnaire survey, which consisted of stress and creativity items.

Creativity Task

For toy generation, the subjects were asked to sketch and label three novel ideas within 10 minutes. They were told, "Imagine that you are employed by a toy company that is in need of new ideas for toys. Within 10 minutes, draw three different toys of your own creative designs. Duplication of toys that currently exist or have already existed is not permitted." This script was quoted from Smith et al. [25].

Stress Manipulation

This experiment used two stress manipulations (threat of shock and performance feedback) on the stress group. The two stress manipulations had been used by Bogdan and Pizzagalli [4]. In case of the threat-of-shock manipulation (no shock was ever actually delivered), the subjects were informed of the possibility that they would receive an "unpleasant but not painful" electrical shock through the electrodes attached to their body. In other words, the subjects were instructed that the possibility they would receive a shock was dependent upon their performance in comparison to past subjects. In the performance-feedback manipulation, the subjects were told that they must be lacking in creativity and less creative than previous participants. These manipulations were implemented according to a fixed pattern, independent of actual performance.

Questionnaire Survey

In order to compare manipulated stress in experiment with perceived stress, we selected PSS as survey items [6, 7]. The PSS measures the degree to which situations are considered stressful, by addressing events experienced beforehand. It was designed to quantify how unpredictable, uncontrollable, and overloaded adults find their lives. We conducted another survey to see whether or not self-reported creativity [20, 23] would agree with actual creative performance. Each item in our survey was measured on a seven-point Likert scale, with answers ranging from "strongly disagree" to "strongly agree." The items in the survey were developed by adopting existing measures validated by other researchers. Self-reported creative performance was measured with three items developed by Oldham and Cummings [20]. The items ($\alpha = .928$) asked subjects to rate the level of creativity and originality in the work that they produced. On the other hand, perceived stress was measured with 10 items ($\alpha = .863$).

3.3 Statistics

For physiological signals and creativity assessment, the differences between the stress group and the non-stress group were analyzed with the Mann-Whitney U Test. This test was performed to test the null hypothesis that the stress group is not different from the non-stress group. The results from the Mann-Whitney U Test are presented with the P-value. Statistical significance was assumed for $P < 0.05$. We investigated the two types of creativity ratings for each group with the Wilcoxon signed ranks test. Finally, we examined the relationship between stress and creativity through descriptive statistics.

4 Results

4.1 Differences between Stress Group and Non-stress Group

Physiological Signals and Creativity Assessment

The relationship between manipulated stress and physiological signals (Normalized Δ GSR, Δ SDNN, and Δ LF/HF ratio) was investigated using the Mann-Whitney U Test. This test is one of the most powerful nonparametric tests, and it is a most useful alternative to the parametric test when the researcher wishes to avoid the t test's assumptions or when the sample sizes are relatively small [24]. Though there was no significant difference between the stress and non-stress groups for normalized Δ GSR and Δ SDNN as shown in Table 1, we confirmed that the stress group had a higher Δ LF/HF ratio value than the non-stress group, with statistical significance. On the other hand, creativity assessed by experts did not have a statistically significant result.

Table 1. Mann-Whitney U Test Results for Physiological Signals and Creativity Assessment

Group	N	Normalized Δ GSR		Normalized Δ SDNN		Normalized Δ LF/HF ratio		Creativity Assessed by Experts	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Stress	16	0.006	0.164	0.022	0.313	1.176	1.706	3.410	0.734
Non-stress	15	-0.043	0.233	0.237	0.814	0.262	0.845	3.081	0.862
Total	31	-0.018	0.199	0.126	0.608	0.734	1.416	3.251	0.803
Two-Tailed Probability		0.540		0.678		0.036 *		0.373	

* Statistically significant at $p < 0.05$.

Creativity Assessed by Experts

Subjects' creativity tasks were assessed by three experts who have published papers on creativity. Assessment criteria were the components (creative, original, and novel) of self-reported creative performance developed by Oldham and Cummings [20]. We tested for whether or not assessments were appropriate using the within-rater

agreement statistic (r_{wg} [12]) and used intraclass correlation coefficients, ICC (1) and ICC (K), so as to estimate raters' similarity [18]. The mean r_{wg} of .933 indicated a high level of agreement among judges on rating the internal team environment; the ICC (1) of .820 demonstrated that raters did not show significant variance; and the ICC (K) of .932 suggested that assessments were reliable.

Self-reported Stress and creative performance

Statistically significant differences between the two groups were observed for perceived stress, not for self-reported creative performance. This result shows that our manipulation of stress was well controlled in the experiment and discriminates the stress group from the other.

Table 2. Mann-Whitney U Test Results for Self-reported Stress and Creative Performance

Group	N	Perceived Stress		Self-reported creative performance	
		Mean	SD	Mean	SD
Stress	16	4.388	1.230	2.708	1.134
Non-stress	15	3.340	0.912	3.244	1.630
Total	31	3.881	1.195	2.968	1.399
Two-Tailed Probability		0.016*		0.349	

* Statistically significant at $p < 0.05$.

Table 3. Wilcoxon Signed Ranks Test of Creativity for Each Group

Stress Group		N	Mean Rank	Sum of Ranks
Negative Ranks	(SC < EC)	12	9.208	110.5
Positive Ranks	(SC > EC)	4	6.375	25.5
Ties	(SC = EC)	0		
Total		16		

Z = -2.198

Two-Tailed Probability = **0.028***

SC: Self-reported creative performance, EC: Creativity assessed by experts.

Non-stress Group		N	Mean Rank	Sum of Ranks
Negative Ranks	(SC < EC)	8	6.625	53
Positive Ranks	(SC > EC)	7	9.571	67
Ties	(SC = EC)	0		
Total		15		

Z = -0.398

Two-Tailed Probability = 0.691

SC: Self-reported creative performance, EC: Creativity assessed by experts

* Statistically significant at $p < 0.05$.

Table 4. Descriptive Statistics for creativity

	Group	Mean	Median	SD	Max	Min
Creativity assessed by experts	Stress	3.410	3.333	.734	4.556	2.222
	Non	3.081	3.222	.862	4.556	1.444
Self-reported creative performance	Stress	2.708	2.833	1.134	4.333	1.000
	Non	3.244	3.333	1.630	6.000	1.000

Table 5. Stem-and-Leaf Plot for Creativity

Creativity assessed by experts				Self-reported creative performance			
Stress Group		Non-Stress Group		Stress Group		Non-Stress Group	
Freq- uency	Stem & Leaf	Freq- uency	Stem & Leaf	Freq- uency	Stem & Leaf	Freq- uency	Stem & Leaf
2	2 . 24	2	1 . 48	3	1 . 003	2	1 . 00
3	2 . 688	4	2 . 2557	1	1 . 6	5	2 . 00000
3	3 . 001	7	3 . 223368	4	2 . 0003	1	3 . 3
3	3 . 566	2	4 . 35	1	2 . 6	4	4 . 0036
4	4 . 0134			3	3 . 003	2	5 . 03
1	4 . 5			1	3 . 6	1	6 . 0
				3	4 . 033		
16		15		16		15	

(Stem width: 1, Each leaf: 1 case)

4.2 Relationship between Stress and Creativity

The two types of creativity ratings for each group

While we did not verify the difference between the two groups for creativity through the Mann-Whitney U Test, we made sure that the stress group has more perceived stress than the other as shown in Table 2. In this view, the subjects are thought to be properly divided. Therefore, we would confirm which group could separately explain creativity by comparing creativity assessed by experts and self-reported creativity. For creative performance, Table 3 explains that there is a discrepancy between the assessments of experts and the self-reported ratings within the stress group, according to the Wilcoxon signed ranks test. In the non-stress group, the two ratings for creative performance have no significant difference. Judging from this result, self-reported creativity in the stress group might be influenced by the manipulated stress. In the next session, these points will be discussed.

Descriptive Statistics for two types of creativity ratings

As shown in Table 5 and Figure 1, self-reported creativity in the stress group (mean = 2.708; median = 2.833) is relatively lower than in the non-stress group (mean = 3.244; median = 3.333), while creativity assessed by experts in the stress group (mean = 3.410; median = 3.333) is relatively higher than in the non-stress group (mean = 3.081; median = 3.222).

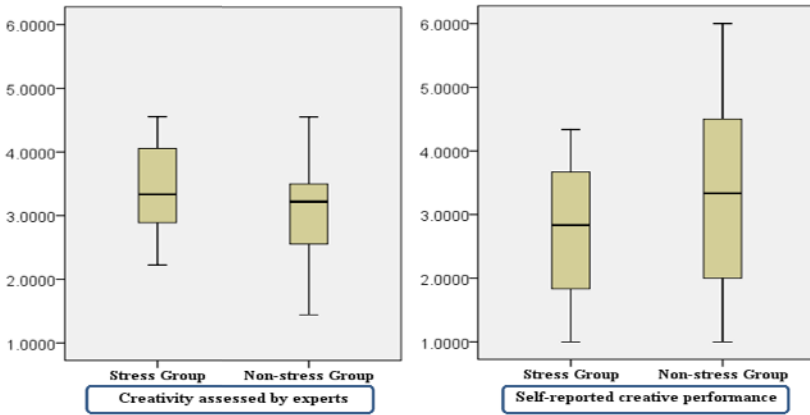


Fig. 1. Box Plot for Creativity

5 Discussion

From mean values of the results, the stress group had a higher normalized Δ GSR, a lower normalized Δ SDNN, and a higher normalized Δ LF/HF ratio than the non-stress group. However, the only normalized Δ LF/HF ratio had statistical significance. Our physiological signals indicated that the stress group has mental stress in the experiment, judging from descriptive statistics. Many published papers have selected Δ LF/HF ratio as an important biomarker of ANS activity associated with mental stress [27]. GSR might not be a critical indicator of mental stress in that sweat-gland activity may be related to ethnic origin [15].

While many researchers have explained the curvilinear relationship of stress and performance, the negative effects of stress become magnified under conditions of acute stress, or stress with sudden, novel, intense, and relatively short duration [22]. Such as threat of shock and negative performance feedback in our experiment, acute stress negatively affects mental models and transactional memory, explaining poorer performance through information-processing theory [11]. For self-reported creative performance, the stress group was evaluated as being a level lower than the non-stress group in the experiment, aside from statistical significance. However, we doubt the inverse relationship of stress and creativity, because the stress group's self-reported creativity might be influenced by negative performance feedback in the experiment.

Judging from experts' assessments, the stress group showed relatively strong creativity. On the other hand, Nicol and Long [19] found that there was no relation between creative thinking and stress for music therapists. In this respect, their studies suggest that stressors' effects on creativity are more complex than previously assumed and point to the need for understanding boundary conditions that shed light on inconsistent findings [19]. For future research, we should investigate other stressors, not negative performance feedback, that may affect self-reported measures. We did not find the relationship of stress and creativity to have statistical significance. We should consider acquiring more sufficient samples for experimentation and including other factors that might influence stress and creativity in order to induce various and delicate interpretations.

Acknowledgments. This study was supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0).

References

1. Amabile, T.M.: *Creativity in Context: Update to the Social Psychology of Creativity*. Westview Press, Boulder (1996)
2. Avey, J.B., Luthans, F., Hannah, S.T., Sweetman, D., Peterson, C.: Impact of Employees' Character Strengths of Wisdom on Stress and Creative Performance. *Human Resource Management Journal* (2011), doi:10.1111/j.1748-8583.2010.00157.x
3. Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., Sleight, P.: Effects of Controlled Breathing, Mental Activity, and Mental Stress with or without Verbalization on Heart Rate Variability. *Journal of the American College of Cardiology* 35(6), 1462–1469 (2000)
4. Bogdan, R., Pizzagalli, D.A.: Acute Stress Reduces Reward Responsiveness: Implications for Depression. *Biological Psychiatry* 60(10), 1147–1154 (2006)
5. Byron, K., Khazanchi, S., Nazarian, D.: The relationship between Stressors and Creativity: A Meta-Analysis Examining Competing Theoretical Models. *Journal of Applied Psychology* 95(1), 201–212 (2010)
6. Cohen, S., Williamson, G.: Perceived Stress in a Probability Sample of the United States. In: Spacapan, S., Oskamp, S. (eds.) *The Social Psychology of Health*. Sage, Newbury Park (1988)
7. Cohen, S., Kamarck, T., Mermelstein, R.: A global Measure of Perceived Stress. *Journal of Health and Social Behavior* 24, 386–396 (1983)
8. Cohen, S., Kessler, R., Gordon, L.: *Measuring Stress - A Guide for Health and Social Scientists*. Oxford University Press (1997)
9. Darrow, C.: The Rationale for Treating the Change in Galvanic Skin Response as a Change in Conductance. *Psychophysiology* 1(1), 31–38 (1964)
10. Dise-Lewis, J.E.: The Life Events and Coping Inventory: An Assessment of Stress in Children. *Psychosomatic Medicine* 50, 484–499 (1988)
11. Ellis, A.P.J.: System Breakdown: the Role of Mental Models and Transactive Memory in the Relationship between Acute Stress and Team Performance. *Academy of Management Journal* 49(3), 576–589 (2006)

12. James, L.R., Demaree, R.G., Wolf, G.: rWG: An Assessment of Within-group Interrater Agreement. *Journal of Applied Psychology* 78, 306–309 (1993)
13. Jex, S.M., Beehr, T.A., Roberts, C.K.: The Meaning of Occupational Stress Items to Survey Respondents. *Journal of Applied Psychology* 77(5), 623–628 (1992)
14. Koh, K., Park, J., Kim, C., Cho, S.: Development of the Stress Response Inventory and its Application in Clinical Practice. *Psychosomatic Medicine* 63, 668–678 (2001)
15. Kugelmass, S., Lieblich, I.: Effects of Realistic Stress and Procedural Interference in Experimental Lie Detection. *Journal of Applied Psychology* 50(3), 211–216 (1966)
16. LePine, J.A., Podsakoff, N.P., LePine, M.A.: A Meta-analytic Test of the Challenge Stressor-hindrance Stressor Framework: An Explanation for Inconsistent Relationships among Stressors and Performance. *Academy of Management Journal* 48, 764–775 (2005)
17. Marsh, R.L., Landau, J.D., Hicks, J.L.: How Examples May (and May Not) Constrain Creativity. *Memory and Cognition* 24(5), 669–680 (1996)
18. McGraw, K.O., Wong, S.P.: Forming Inferences about Some Intraclass Correlation Coefficients. *Psychological Methods* 1, 30–46 (1996)
19. Nicol, J.J., Long, B.C.: Creativity and Perceived Stress of Female Music Therapists and Hobbyists. *Creativity Research Journal* 9(1), 1–10 (1996)
20. Oldham, G.R., Cummings, A.: Employee Creativity: Personal and Contextual Factors at Work. *Academy of Management Journal* 39, 607–634 (1996)
21. Ritvanen, T., Louhevaara, V., Helin, P., Vaisanen, S., Hanninen, O.: Responses of the Autonomic Nervous System during Periods of Perceived High and Low Work Stress in Younger and Older Female Teachers. *Applied Ergonomics* 37, 311–318 (2005)
22. Salas, E., Driskell, J.E., Hughes, S.: Introduction: the Study of Stress and Human Performance. In: Driskell, J.E., Salas, E. (eds.) *Stress and Human Performance*. Erlbaum, Mahwah (1996)
23. Shalley, C.E., Gilson, L.L., Blum, T.C.: Interactive Effects of Growth Need Strength, Self-reported Creative Performance. *Academy of Management Journal* 52(3), 489–505 (2009)
24. Siegel, S., Castellan, N.J.: *Nonparametric Statistics for the Behavioural Sciences*. McGraw-Hill, New York (1988)
25. Smith, S.M., Ward, T.B., Schumacher, J.S.: Constraining Effects of Examples in a Creative Generation Task. *Memory & Cognition* 21(6), 837–845 (1993)
26. Van der Kar, L.D., Blair, M.L.: Forebrain Pathways Mediating Stress Induced Hormone Secretion. *Frontiers in Neuroendocrinology* 20, 41–48 (1999)
27. Zhong, X., Hilton, H.J., Gates, G.J., Jelic, S., Stern, Y., Bartels, M.N., DeMeersman, R.E., Basner, R.C.: Increased Sympathetic and Decreased Parasympathetic Cardiovascular Modulation in Normal Humans with Acute Sleep Deprivation. *Journal of Applied Physiology* 98(6), 2024–2032 (2005)

The Structure of Individual Creativity Revelation Processes with Task Characteristics and Social Characteristics: An Empirical Analysis Based on IT-Services Participants

Do Young Choi¹, Min Hee Hahn², and Kun Chang Lee^{3,*}

¹ Principal Consultant, Solution Business Unit, LG CNS Co., Ltd,
Seoul 100-725, Republic of Korea

² SKK Business School, Sungkyunkwan University,
Seoul 110-745, Republic of Korea

³ Professor at SKK Business School, WCU Professor at Department of Interaction Science,
Sungkyunkwan University, Seoul 110-745, Republic of Korea
{dychoi96,minheehahn,kunchanglee}@gmail.com

Abstract. Creativity is considered one of the most important strategic resources in knowledge-based industry. This research proposes an individual-creativity-revelation model in the business environment, and mainly focuses on the relations among task characteristics, social characteristics, and creativity-revelation processes. Based on a sample of 520 employees in the information-technology-services industry, this research investigate the effects of task characteristics – task complexity and task stress – and social characteristics – trust and social capital – on creativity-revelation processes on an individual level. We adopted two concepts of creativity-revelation processes – exploitation and exploration – and the results indicate that task characteristics and social characteristics have relations with the creativity-revelation processes.

Keywords: Creativity, Exploitation, Exploration, Task Complexity, Task Stress, Trust, Social Capital.

1 Introduction

Creativity has been considered one of the most important factors for group-level or organizational-level performances. Also, business organizations have focused on appropriate creativity management for enhancing value creation in the environment of rapid market change. Therefore, many researchers have studied the influencing factors on creativity and the creativity-revelation mechanism in the business environment. Researchers expanded the influencing factors on creativity and their relations from the individual level to the group level and organizational level [8, 22]. When we consider that individuals typically perform their tasks with other members of organizations, the characteristics of the working environment and social environment should be considered simultaneously beside the characteristics of the

* Corresponding author.

individual level in order to address their effects on creativity and in order to explain creativity revelation. For example, social relations, organizational atmosphere, and task complexity could be considered as environmental and social factors affecting creativity. Also, intellectual capability, self-efficacy, and knowledge could be considered as individual characteristics affecting creativity [22]. Meanwhile, creation and generation of new ideas and new knowledge has its revelation processes. Researchers have focused on exploitation and exploration as processes affecting innovation and creativity [19].

This research addresses the relations between environmental factors and creativity-revelation processes among several influencing factors on individual creativity. Specifically, we focus on how environmental factors can affect creativity-revelation processes for individual creativity. Although several factors among environmental characteristics could be considered to address individual creativity, we will focus on two perspectives of environmental characteristics – task perspectives, including task complexity and task stress, and social perspectives, including trust and social capital. In order to explain these research questions, we conducted a questionnaire survey on information-technology (IT) professionals at IT-services companies in Korea. And this research used the structural-equation model and analyzed the relations among influencing factors and creativity-revelation processes.

2 Theoretical Foundations and Hypotheses

2.1 Creativity, Exploitation, and Exploration

Creativity has been defined differently in various fields and typically it can be defined as processes that can induce innovative results by creating something new and innovative [2]. Researchers have expanded the coverage of the creativity-research area from the individual level to the group or organizational level [8, 22]. Regarding influencing factors on individual creativity, Woodman et al. [22] proposed that cognitive awareness, personality, intrinsic motivation, and knowledge could be regarded as influencing factors in the individual level. Furthermore, they explained influencing factors of group level and organizational level such as organizational cohesiveness, diversity, culture, resources, and structure. The notable point of their research was that the factors based on three levels – individual, group, and organization – could affect one another and consequently creativity could be revealed by interactions. Meanwhile, researchers found that interactions in organizations could affect creativity [12]. Specifically, exploitation and exploration are considered important processes affecting innovation, because new-idea and knowledge generation may come from creativity [19]. Therefore, this research focuses on the relationship between individual creativity and creativity-revelation processes – exploitation and exploration.

2.2 Task Characteristics and Social Characteristics

Research on influencing factors on individual creativity has considered several perspectives – individual characteristics, task characteristics, social factors such as

working environment and organizational atmosphere, etc. This research focused on task complexity and task stress among several factors of task characteristics. Also, trust and social capital were considered as influencing factors on individual creativity among social characteristics. Regarding task complexity, much of the research addressed how several factors of individual characteristics such as cognitive characteristics, problem-solving styles, leadership style, and task complexity could affect individual creativity [3, 15, 16]. Meanwhile, the results of task stress were found to fall within three types: behavioral result, cognitive result, and physical result. Among these three results, the cognitive result is the most strongly related with creativity, which reduces concentration power and decision capability [18].

Among the social characteristics this research focuses on, trust is the most fundamental factor for efficient communication. With it, members of organizations can voluntarily recognize and accept each other's authority and role [21]. Moreover, it could enhance collaboration behavior and cooperation [4]. According to McCauley and Kuhnert [11], trust among members within an organization is a kind of environmental factor strongly related with innovative behavior. Therefore, trust has strong relations with creativity. Meanwhile, Tsai and Ghoshal [20] consider social capital as an attribute of organization. They insist that members of an organization could know their organization's goal and the appropriate knowledge and behavior necessary for achieving its goal when social capital is implemented within the organization. Consequently, they could create new value continuously by collaborating and supporting each other and exchanging and combining existing resources.

2.3 Research Model and Hypotheses

This research intends to investigate the relations among task characteristics and social characteristics and their effects on individual creativity processes. In that sense, Figure 1 shows the research model, which consists of seven constructs – task complexity, task stress, trust, social capital, exploitation, exploration, and individual creativity. This model assumed that task complexity, trust, and social capital could have positive effects on individual creativity through the mediation of creativity processes – exploitation and exploration. Also, task stress was assumed to have a negative effect on individual creativity through the same mediation of creativity processes.

What tasks are complex means that team members are confronted with unexpected status for problem solving [1]. Also, team members with high task complexity need new knowledge and creative solutions for solving vague problems [1]. That is, if the tasks that team members face are complex, then the tasks need more effort from the individuals in order to generate or create new knowledge, new ideas, and new solutions. Meanwhile, exploitation is defined as an innovative process to create new knowledge or solutions based on existing knowledge, and exploration is defined as a process of new knowledge or solution creation without regard to existing knowledge [9]. It could be interpreted that team members should allocate more efforts to creative

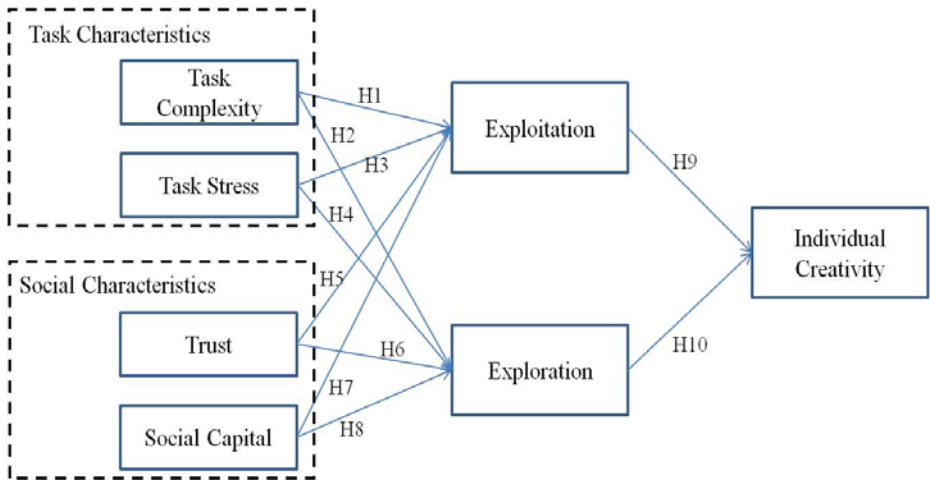


Fig. 1. Individual Creativity Revelation Model

processes – exploitation and exploration - for generating new solutions in order to solve complex tasks. Therefore, we have derived the following research hypotheses with regard to the relations between task complexity and creativity processes – exploitation and exploration.

- H1.* Task complexity influences exploitation positively.
H2. Task complexity influences exploration positively.

It is considered that task stress has a negative effect on cognitive results, which could reduce concentration and decision capability [18]. Therefore, it could be inferred that team members with high task stress could not concentrate on creativity processes.

- H3.* Task stress influences exploitation negatively.
H4. Task stress influences exploration positively.

When the level of trust is high among team members, members can voluntarily recognize and accept each member's authority and role [21]. And, trust could expedite collaboration among members [4]. According to the research of McCauley and Kuhnert [11], trust is strongly related with innovative behavior. Consequently, we can infer that teams with a high level of trust have strong foundations for creativity and its processes. Therefore, we have derived the following research hypotheses with regard to the relations between trust and creativity processes.

- H5.* Trust influences exploitation positively.
H6. Trust influences exploration positively.

Tsai and Ghoshal [20] insist that social capital could create new value continuously by collaborating and supporting each other and exchanging and combining existing resources. When we consider that exploitation and exploration could be expedited by collaboration and knowledge sharing among team members, we can derive the following two research hypotheses.

- H7.* Social capital influences exploitation positively.
- H8.* Social capital influences exploration positively.

Exploitation and exploration are both considered key processes for innovation and organizational adaptation [9, 10]. When we consider that innovative activity and innovation processes have a strong relationship with creativity [7, 15], we can derive the following research hypotheses with regard to the relations between creativity processes and individual creativity.

- H9.* Exploitation influences individual creativity positively.
- H10.* Exploration influences individual creativity positively.

3 Research Methodology and Analysis

3.1 Sample and Data Collection

Questionnaire surveys were conducted to investigate the individual-creativity-revelation model this research proposed. This research adapted proven measurement items from reliable previous studies and all measurement items were measured on 7-point Likert-scale. For example, we adapted four measurement items from the research done by Ettlie and O'Keefe [6] and Munoz-Doyague et al. [13] in the case of individual creativity. In order to measure task complexity and task stress, seven measurement items were adapted from Daft and Macintosh [5], and eight survey items from Sosik and Godshalk [17]. Also, measurement items for exploitation and exploration were adapted from Prieto et al. [14]. After building the survey questionnaire, we collected 560 questionnaires from IT professionals at IT-services companies in Korea. Among responses, 520 responses were appropriate for the analysis of this research.

3.2 Measures

PLS (Partial Least Squares) analysis was conducted in order to verify our research model and in order to test hypotheses by using SmartPLS 2.0 software. PLS is considered a useful and appropriate method in the areas of structural equation modeling of which the theory and variables' relations are weak.

Reliability and validity analyses were conducted in order to verify consistency of the measurement items with our intent. The values of Cronbach alpha of all constructs

are higher than 0.7, which indicates they have a high internal consistency. First the t-test for factor loading was reviewed in order to assess convergent validity. Also, convergent and discriminant validity was assessed in order to examine composite reliability and AVE (Average Variance Extract) as shown in Table 1. Table 2 shows the correlation matrix among constructs and the square root of AVE on the diagonal; the discriminant validity was assured from the analysis.

Table 2. Reliability and Convergent Validity

Construct	Measurement Item	Factor Loading	Cronbach's α	Composite Reliability	AVE
Creativity	CR1	0.8481	0.8704	0.9118	0.7216
	CR2	0.7805			
	CR3	0.8762			
	CR4	0.8888			
Exploitation	ET1	0.8264	0.8857	0.9134	0.6381
	ET2	0.8409			
	ET3	0.8319			
	ET4	0.8069			
	ET5	0.7094			
	ET6	0.7692			
Exploration	ER1	0.8541	0.8942	0.9265	0.7592
	ER2	0.8873			
	ER3	0.8808			
	ER4	0.8626			
Social Capital	SC1	0.8449	0.9142	0.9358	0.7447
	SC2	0.8751			
	SC3	0.8523			
	SC4	0.8493			
	SC5	0.8923			
Task Complexity	TC3	0.7773	0.8539	0.8958	0.6329
	TC4	0.8084			
	TC5	0.8329			
	TC6	0.8379			
	TC7	0.7149			
Task Stress	ST1	0.7798	0.9137	0.9322	0.7338
	ST2	0.8830			
	ST3	0.8415			
	ST7	0.9164			
	ST8	0.8564			
Trust	TR1	0.7496	0.9433	0.9531	0.7185
	TR2	0.7546			
	TR3	0.8671			
	TR4	0.8685			
	TR5	0.8581			
	TR6	0.8719			
	TR7	0.9010			
	TR8	0.8958			

Table 2. Discriminant Validity

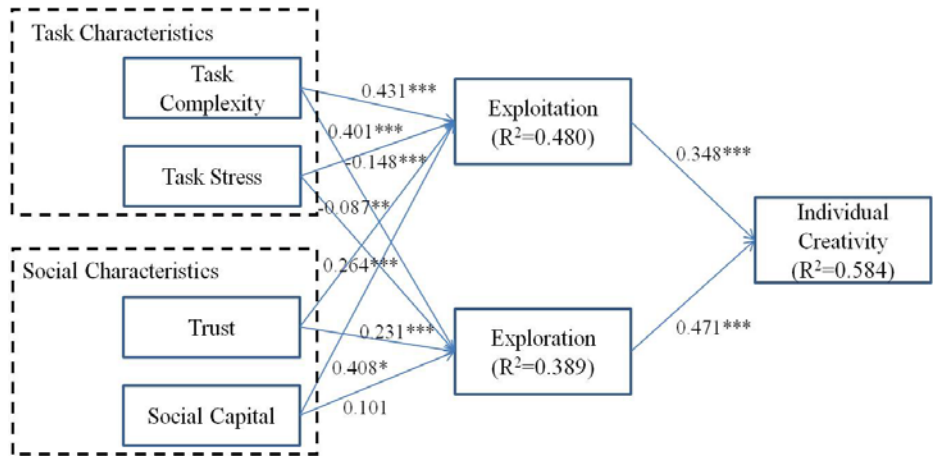
Construct	CR	ET	ER	SC	TC	TS	TR
CR	0.8495						
ET	0.6934	0.7988					
ER	0.7264	0.7329	0.8713				
SC	0.4315	0.4837	0.4410	0.8630			
TC	0.5401	0.6014	0.5537	0.4338	0.7956		
TS	-0.1265	-0.1345	-0.0740	0.0453	0.0455	0.8566	
TR	0.4675	0.5622	0.5057	0.7388	0.4916	-0.0415	0.8476

CR: Creativity, ET: Exploitation, ER: Exploration, SC: Social Capital, TC: Task Complexity, TS: Task Stress, TR: Trust.

3.3 Analysis and Results

The hypotheses of the model this research proposed were tested by coefficients of path and R² as shown in Figure 2. The coefficients of path show how strong the relations are among respondent variables and explanatory variables. Also, the value of R² indicates the amount of variance explained by the explanatory variables.

The model shows that the R² values of all constructs are higher than the 10% that is recommended. This result strongly supports the posited relationships among constructs. The R² value of individual creativity, which is the last respondent variable explained by all independent variables, is 58.4%, that of exploitation is 48.0%, and that of exploration is 38.9%. Table 3 summarizes the result of the research hypothesis. As shown in the Table 3, hypothesis 8 was not accepted. But, other hypotheses were all accepted.



* p<0.05, ** p<0.01, *** p<0.001.

Fig. 2. Research Results

Table 3. Results of Hypotheses Testing

H-No.	Path Name	Path Coefficient	t-Value	Result
H1	TC → ET	0.431	10.767***	Accept
H2	TC → ER	0.401	7.936***	Accept
H3	TS → ET	-0.148	3.465***	Accept
H4	TS → ER	-0.087	2.014***	Accept
H5	TR → ET	0.264	4.655***	Accept
H6	TR → ER	0.231	3.917***	Accept
H7	SC → ET	0.408	1.806***	Accept
H8	SC → ER	0.101	1.523	Reject
H9	ET → CR	0.348	7.104***	Accept
H10	ER → CR	0.471	10.038***	Accept

CR: Creativity, ET: Exploitation, ER: Exploration, SC: Social Capital, TC: Task Complexity, TS: Task Stress, TR: Trust.

*p<0.05, **p<0.01, ***p<0.001.

4 Discussion and Concluding Remarks

This research has proposed the individual-creativity-revelation model in the business environment, which mainly focuses on the relations among task characteristics, social characteristics, and creativity revelation processes. Based on the survey of IT professionals in the IT-services industry, this research investigated the effects of task characteristics – task complexity and task stress – and social characteristics – trust and social capital – on creativity-revelation processes on an individual level. Also, in order to find out the effect on individual creativity, we adopted two concepts of creativity-revelation processes – exploitation and exploration.

The result of PLS analysis shows that task characteristics, which consisted of task complexity and task stress, and social characteristics, which consisted of trust and social capital, have significant effects on the creativity revelation processes – exploitation and exploration. Moreover, it could be found that both creativity-revelation processes – exploitation and exploration – have significant effects on individual creativity. Therefore, we can infer that creativity-revelation processes and individual creativity could be reinforced by task characteristics and social characteristics.

Based on the findings from this empirical research, there are several implications for creativity management. Regarding task characteristics, we found out that task complexity had a positive effect on creativity processes and individual creativity. Relatively routine jobs and tasks might not need creativity. But, task complexity encourages team members to seek new solutions and new knowledge through creative processes. Therefore, managers should pay attention to task complexity for creativity improvement, which means tasks should be relatively complex when creativity is

needed. Task stress had a negative effect on creativity processes and individual creativity, as expected. Managers should consider that creativity could be maximized when team members have less-stressful working situations. Secondly, regarding social characteristics, it was found that both trust and social capital had positive effects on creativity processes – exploitation and exploration – and individual creativity. Practically, managers should encourage a favorable working environment, where trust and social capital could be implemented within teams. Consequently, team members willingly exchange their knowledge and collaborate to seek new knowledge and new solutions for problems they meet.

Nevertheless, there are several limitations and considerations for further study. Firstly, this research considers only four factors for addressing creativity-revelation processes and individual creativity. There exist many influencing factors on individual creativity and several considerable variables on creativity-revelation processes. Other factors should be included in the future study, such as psychological characteristics and social relations – social network structure – etc. Secondly, this research involved a survey targeting IT professionals in the IT-services industry. This result might be considered too limited to be generalized. Therefore, future study needs to extend the scope for surveys.

Acknowledgment. This research was supported by WCU(World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0).

References

1. Akgun, A.E., Byrne, J., Keskin, H., Lynn, G.S., Imamoglu, S.Z.: Knowledge networks in new product development projects: A transactive memory perspective. *Information & Management* 42(8), 1105–1120 (2005)
2. Amabile, T.M.: A model of creativity and innovation in organizations. *Research in Organizational Behavior* 10, 123–167 (1988)
3. Baer, M., Oldham, C.R., Cummings, A.: Rewarding creativity: When does it really matter? *Leadership Quarterly* 14, 569–586 (2003)
4. Burt, R.S., Knez, M.: Trust and third party gossip. In: Kramer, R.M., Tyler, T.R. (eds.) *Trust in Organization: Frontiers of Theory and Research*, pp. 68–89 (1996)
5. Daft, R.L., Macintosh, N.B.: A tentative exploration into the amount and equivocality of information processing in organizational work units. *Administrative Science Quarterly* 26(2), 207–224 (1981)
6. Ettlie, J.E., O’Keefe, R.D.: Innovative attitudes, values, and intentions in organizations. *Journal of Management Studies* 19(2) (1982)
7. Kanter, R.M.: When a thousand flowers bloom: structural, collective, and social conditions for innovation on organization. *Research in Organizational Behavior* 10, 169–211 (1988)
8. Kurtzberg, T.R., Amabile, T.M.: From Guilford to creative synergy: opening the balck box of team-level creativity. *Creativity Research Journal* 13, 285–294 (2001)
9. Lazer, D., Friedman, A.: The network structure of exploration and exploitation. *Administrative Science Quarterly* 52, 667–694 (2007)

10. March, J.G.: Exploration and exploitation in organizational learning. *Organization Science* 2(1), 71–87 (1991)
11. McCauley, D.P., Kuhnert, K.W.: A theoretical review and empirical investigation of employee trust in management. *PAQ*, 265–287 (Summer 1992)
12. Milgram, R.M., Rabkin, L.: Developmental test of Mednick's associative hierarchies of original thinking. *Developmental Psychology* 16, 157–158 (1980)
13. Munoz-Doyague, M.F., Gonzalez-Alvarez, N., Nieto, M.: An examination of individual factors and employees' creativity: The case of Spain. *Creativity Research Journal* 20(1), 21–33 (2008)
14. Prieto, I.M., Revilla, E., Rodriguez-Prado, E.: Managing the knowledge paradox in product development. *Journal of Knowledge Management* 13, 157–170 (2009)
15. Scott, S.G., Bruce, R.: Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal* 37, 580–607 (1994)
16. Shin, S., Zhou, J.: Transformational leadership, conservation, and creativity: Evidence from Korea. *Academy of Management Journal* 46, 703–714 (2003)
17. Sosik, J.J., Godshalk, V.M.: Leadership styles, mentoring functions received, and job-related stress: a conceptual model and preliminary study. *Journal of Organizational Behavior* 21, 365–390 (2000)
18. Spector, P.E.: Perceived control by employees: A meta-analysis of studies concerning autonomy and participation at work. *Human Relations* 39(11), 1005–1017 (1986)
19. Taylor, A., Greve, H.R.: Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal* 49(4), 723–740 (2006)
20. Tsai, W., Ghoshal, S.: Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal* 40(4), 464–476 (1998)
21. Tyler, T.R., DeGoey, P.: Trust in organizational authorities: The influence of motive attributions on willingness to accept decisions. In: Kramer, R.M., Tyler, T.R. (eds.) *Trust in Organization: Frontiers of Theory and Research*, pp. 331–356. Sage Publications (1996)
22. Woodman, R.W., Sawyer, J.E., Griffin, R.W.: Toward a theory of organizational creativity. *Academy of Management Review* 18(2), 293–321 (1993)

Effects of Social and Emotional Intelligence on the Creative Process and Individual Creativity

Min Hee Hahn¹, Do Young Choi², and Kun Chang Lee^{3,*}

¹SKK Business School, Sungkyunkwan University,
Seoul 110-745, Republic of Korea

²Principal Consultant, Solution Business Unit, LG CNS Co., Ltd,
Seoul 100-725, Republic of Korea

³Professor at SKK Business School, WCU Professor at Department of Interaction Science
Sungkyunkwan University, Seoul 110-745, Republic of Korea
{minheehahn, dychoi96, kunchanglee}@gmail.com

Abstract. This study proposes an individual creativity model that consists of personal psychological characteristics and creative processes. We argue that social intelligence and emotional intelligence, as personal psychological characteristics, significantly influence creativity, and analyze how they are related to the creative process and individual creativity. We assume that the creative process includes both exploitation and exploration, where exploitation is considered an existing use of the solution and exploration is believed to aid in the development of new solutions. Using a structural equation model to analyze 447 valid questionnaires collected from employees in the Korean IT industry, we find that the creative process and individual creativity are reinforced by high levels of social and emotional intelligence. Another interesting finding is that exploration reinforces individual creativity, while exploitation does not directly strengthen individual creativity. *Keywords:* Creativity, Exploration, Exploitation, Social intelligence, Emotional intelligence.

1 Introduction

Since Guilford [10] stressed the importance of creativity in the American Psychology Association (APA), related research has been carried out based on various approaches. The importance of social intelligence, which makes human relationships more flexible and improves job performance, is increasing, and researchers such as Goleman [9] have argued that complex ‘relationships’ are very important in achieving social success. Goleman also argues that emotional intelligence and related practical and creative intelligence will lead the 21st century. As such, social and emotional intelligence are drawing the focus of academia and the business market as personal psychological characteristics.

This research aims to show how social and emotional intelligence (the new paradigm as personal psychological variable factors) affect individual creativity. We study how exploration, or searching for new knowledge, and exploitation, or using

* Corresponding author.

existing knowledge, are related to personal psychological characteristics and individual creativity. First, we investigate the multidimensional relationship between social intelligence and emotional intelligence as psychological characteristics variables, exploitation and exploration as part of the creative process, and individual creativity. Second, by carrying out an empirical analysis of the relationship between the creative process, individual creativity, and the personal psychological characteristics for the improvement of individual creativity, we provide implications from both academic and practical perspectives.

2 Previous Studies

2.1 Social and Emotional Intelligence

(1) Social intelligence

Social intelligence means the ability to behave reasonably based on understanding the emotions, thoughts, and behavior of ourselves and other people in everyday life. As Guilford [11] suggested regarding the structure of the intellect model, more research in this field has been conducted in recent years. Goleman [9] recently defined social intelligence as the ability to lead relationships with others positively and effectively, making them harmonious and productive in the border of society.

In 1980, research on the multidimensional characteristics of social intelligence was first conducted [9]. For example, Wong *et al.* [20] discussed social perception, social insight, and social knowledge as components of social intelligence. Various components of social intelligence mentioned in the theory of Thomdike [19] have recently been developed and include the study of consensus and interaction in society by Goleman [9]. In this research, we also measure social intelligence based on his research.

(2) Emotional intelligence

The first researchers to use the concept of emotional intelligence are Salovey and Mayer [17], whose work has received increasing attention since they defined emotional intelligence as “*the subset of social intelligence that involves the ability to monitor one’s own and others’ feelings and emotions, to discriminate among them and to use this information to guide one’s thinking and actions.*” Goleman later introduced the term “emotional intelligence” in *The Times* in 1995, and the term has since received a great attention from researchers around the world [4].

Wong and Law [21] developed a new EI scale following the four-dimensional definition of EI introduced by Davies *et al.* [5]. Wong and Law insisted that emotional intelligence is composed of the ability to clearly understand and express self-emotion (*Self-Emotions Appraisal (SEA)*), the ability to recognize and understand the emotions of others (*Others-Emotions Appraisal (OEA)*), the ability to use the emotional information that individual has in achieving greater performance (*Use of Emotion (UOE)*), and the ability to control personal emotion in accordance with the given situation (*Regulation of Emotion (ROE)*). In this research, we measured emotional intelligence using the results of Wong and Law.

2.2 Exploitation and Exploration

The concept of exploitation and exploration was first introduced by March [15] and is currently regarded as an important mechanism in the growth and survival of an organization. Levinthal and March [14] defined exploration as “*the pursuit of new knowledge, of things that might come to be known*” and exploitation as “*the use and development of things already known.*” In this study, we accepted these concepts of exploitation and exploration and adopted the previous definitions.

Many researchers have carried out studies on exploration-exploitation from various perspectives. For example, some researchers have applied these constructs to strategic alliances, product development, and organizational innovation and performance [12]. Studies on exploration and exploitation have used the two constructs as dependent variables, examining the influence of specific independent variables on them or the effect of diverse independent variables on firm performance and its substitutes. For instance, Nerkar [16] investigated the impacts of temporal exploitation and exploration on later knowledge creation.

In this study, we do not try to put out comprehensive interpretations about these, but focus on the perspective of balancing exploration and exploitation to improve individual creativity. Previous studies indicate that maintaining an appropriate balance between exploration and exploitation activities is a primary factor in a firm's survival and prosperity [15].

2.3 Individual Creativity

The history of creativity research, which started with the speech of Guilford [10], the father of creativity study, at an American Psychology Association (APA) meeting, can be categorized according to different time periods: basic research on creativity in the 1950-60s, research on the individual difference variables and learning program in the 1970-1980s, and lastly, research based on the ecological approach as the dimensional perspective that stresses interactions among the creativity component factors from the 1990s to the present.

Since the 1990s, more research on individual creativity has been carried out in the field of management, and this research can be divided into two sections. The first includes research focused only on the independent variables that affect creativity and the second is focused on the moderate variables. Research focused only on the independent variables is composed of three parts. First is the study of the personal characteristics that affect creativity, including intelligence, cognitive style, and personality. Second are the research studies investigating the contextual characteristics that affect creativity, such as leadership [18], the creative atmosphere of the organization [7], and expectations of coworkers regarding this type of research. Last is research that addresses the interactions among personal and contextual characteristics, and this area of research is actively being carried out. Previous research on individual creativity has mainly focused on personal and contextual characteristics and their interactions.

3 Research Model

Exploitation and exploration have been discussed as part of the process leading to creative or innovative performance. Researchers have suggested that social intelligence and emotional intelligence are important factors that affect creativity [1].

3.1 Personal Psychological Characteristics and Creativity

According to the research of Barchard [2], social intelligence is composed of the skills that create useful social products and is an important factor in solving social problems. In addition, Goleman [9] viewed satisfaction and the sense of responsibility as strong indicators of creative intelligence in dealing with the meaning and value of problems as well as their solutions. After all, individuals with high social intelligence can understand and control the emotions of other people relatively well and improve the performance of an organization through harmonization. Therefore, we propose the following hypotheses to investigate the relationship between social intelligence and individual creativity based on the discussion of Goleman [9], who argued that social intelligence leads to an increase in an individual's creativity index and an improvement in job performance.

Hypothesis 1: Social intelligence positively influences exploitation.

Hypothesis 2: Social intelligence positively influences exploration.

Hypothesis 3: Social intelligence positively influences individual creativity.

Isen [13] studied the relationship between laughing and problem solving that requires creative solutions. He found that stronger emotion takes place when individuals face a problematic situation, and this emotion leads the individual's attention to the new problem. In addition, Goleman [8] discussed emotional intelligence and emotional competency, arguing that learned emotional competency, which improves performance based on emotional intelligence, is two times as important as cognitive ability in predicting staff members' creative performance. Lastly, Fabio [6] found that leaders with the highest performance cause more than three times the amount of laughter than leaders with average performance. This indicates that a more positive atmosphere leads individuals to accept information more effectively and helps them react more sensibly and creatively. Therefore, taking into account the results of preceding researchers, we derived the following research hypotheses.

Hypothesis 4: Emotional intelligence positively influences exploitation.

Hypothesis 5: Emotional intelligence positively influences exploration.

Hypothesis 6: Emotional intelligence positively influences individual creativity.

3.2 Creative Process and Individual Creativity

Exploitation seeks a differentiated solution from the existing solution by using existing knowledge, while exploration is carried out as an innovative activity. A new solution is found in such a process and new knowledge is gained through the learning

process [15]. Audia and Goncalo [1] divided creativity into two concepts: divergent creativity and incremental creativity, explaining divergent creativity and incremental creativity in relation to exploration and exploitation, respectively.

It is natural to assume that exploration and exploitation activities are related to creative activity as well as individual creativity. The theory of exploration-exploitation is potentially useful for understanding the creative process because it incorporates past success as a factor impacting the propensity to explore new ideas [1]. Therefore, we propose the following research hypotheses.

Hypothesis 7: Exploitation positively influences individual creativity.

Hypothesis 8: Exploration positively influences individual creativity.

4 Experiment and Results

4.1 Data Collection

The variables included in this study were measured based on a multi-item scale following a 7-point Likert Scale format (1 strongly disagree, 7 strongly agree). The measurement items of each variable were modified based on the reliability and validity of items found in previous studies. For the empirical analysis, we set the staff members who consult and respond to customers in the Korean IT industry as the target group for the analysis. In September 2011, the online survey was carried out over a period of one week. Surveys with errors or missing responses were excluded from the analysis, for a total of 447 respondents included in the present analysis.

From the analysis, we can see that there are more men than women in the sample (368 men, 82.3%, 79 women, 17.7%). In regards to age, 88 respondents are 19-29 years (19.7%), 276 are 30-39 years (61.7%), 72 are 40-49 years (16.1%), and 11 are 50 years or older (2.5%). For work experience, 180 (40.3%) had worked less than 5 years, 127 (28.4%) had worked between 5 and 10 years, 93 (20.8%) had worked between 10 and 15 years, and 47 (10.5%) had worked more than 15 years. About 60% of respondents had more than 5 years of work experience. The job types include IT consulting (25.1%), IT planning (22.1%), R&D (21.7%), system analyst (13.0%), customer requirement analysis (9.2%), and IT analyst (8.9%). All of the respondents are in a job that requires the ability to find new solutions and use existing knowledge by understanding and fulfilling the requirements of customers.

4.2 Reliability and Construct Validity

To analyze the theoretical research model and our own research hypotheses, we used SmartPLS 2.0, a type of partial least squares (PLS) software. PLS is a structural equation modeling tool that uses a component-based approach for estimation, so it places minimal restrictions on the sample size and residual distribution [3]. Therefore, PLS is especially useful in areas where there is weak theory and limited understanding of the relationships between variables.

Table 1. Reliability and Convergent Validity

Construct		Measurement Item	Factor Loading	Cronbach's α	Composite Reliability	AVE
Social Intelligence		SQ1	0.550	0.850	0.887	0.529
		SQ2	0.627			
		SQ3	0.713			
		SQ4	0.644			
		SQ5	0.691			
		SQ6	0.716			
		SQ7	0.617			
Emotional Intelligence	Self-Emotions Appraisal (SEA)	EQ_SEA1	0.749	0.919	0.930	0.526
		EQ_SEA2	0.809			
		EQ_SEA3	0.792			
		EQ_SEA4	0.549			
	Others-Emotions Appraisal (OEA)	EQ_OEA1	0.644			
		EQ_OEA2	0.807			
		EQ_OEA3	0.815			
		EQ_OEA4	0.747			
	Use of Emotion (UOE)	EQ_UOE1	0.733			
		EQ_UOE2	0.781			
		EQ_UOE3	0.774			
		EQ_UOE4	0.746			
	Regulation of Emotion (ROE)	EQ_ROE1	0.723			
		EQ_ROE2	0.796			
		EQ_ROE3	0.830			
		EQ_ROE4	0.863			
Exploitation		ET1	0.771	0.866	0.901	0.602
		ET2	0.804			
		ET3	0.632			
		ET4	0.588			
		ET5	0.640			
		ET6	0.715			
Exploration		ER1	0.719	0.871	0.912	0.722
		ER2	0.716			
		ER3	0.653			
		ER4	0.607			
Individual Creativity		IC1	0.578	0.936	0.946	0.597
		IC2	0.592			
		IC3	0.560			
		IC4	0.663			
		IC5	0.650			
		IC6	0.589			
		IC7	0.710			
		IC8	0.699			
		IC9	0.735			
		IC10	0.792			
		IC11	0.765			
		IC12	0.758			

* Note : The construct which is composed of 2nd order factor such emotional intelligence are analyzed by directly connecting all indicators used in 1st order factor to 2nd order factor.

We conducted reliability and validity analyses to examine whether the questionnaire items matched our intent. The items were first tested for scale reliability. The Cronbach's alpha scores all exceeded 0.7, indicating high internal consistency. The convergent validity was assessed by reviewing the t-test for factor loading. The convergent validity and discriminant validity were also assessed by examining the composite reliability and average variance extracted (AVE; See Table 1).

As we can see from the factor loading scores presented in Table 1, all scores are above 0.5, indicating that the measurement items do a good job in explaining the constructs. The composite reliability and AVE measures exceed the threshold of 0.7 and 0.5, respectively, and therefore we can conclude that the convergent validity of the measurement model is reasonable.

The discriminant validity of the instrument was assessed by examining the correlations among the questions. For discriminant validity, a measure should correlate with all measures of the same construct more highly than it does with any measures of other constructs. For satisfactory discriminant validity, the average variance extracted (AVE) from the construct should be greater than the variance shared between the construct and other constructs in the model. Table 4 presents the correlation matrix with correlations among the constructs and the square root of AVE on the diagonal. As shown in Table 2, the analysis of the discriminant validity showed that it was at an acceptable level.

Table 2. Discriminant Validity

Construct	EQ	ER	ET	IC	SQ
EQ	0.726				
ER	0.646	0.850			
ET	0.725	0.699	0.776		
IC	0.718	0.747	0.632	0.773	
SQ	0.705	0.629	0.641	0.720	0.728

* Note : SQ (Social Intelligence), EQ (Emotional Intelligence), ET (Exploitation), ER (Exploration), IC (Individual Creativity).

4.3 Hypothesis Testing and Interpretation

The results of the analysis of the causality in the structural equation model are shown in Fig. 1.

The interpretation of the PLS results is based on the R^2 values of the dependent variable, which is explained by the independent variables, as well as through the path coefficient's size, sign, and statistical significance. In this model, the R^2 values of all constructs are above the 10% recommended value. These results provide strong support for the posited relationships among the constructs. The R^2 value of individual creativity, which is the last dependent variable explained by the independent variables, is 69.2%, with R^2 values of 55.9% for exploitation and 47.7% for exploration.

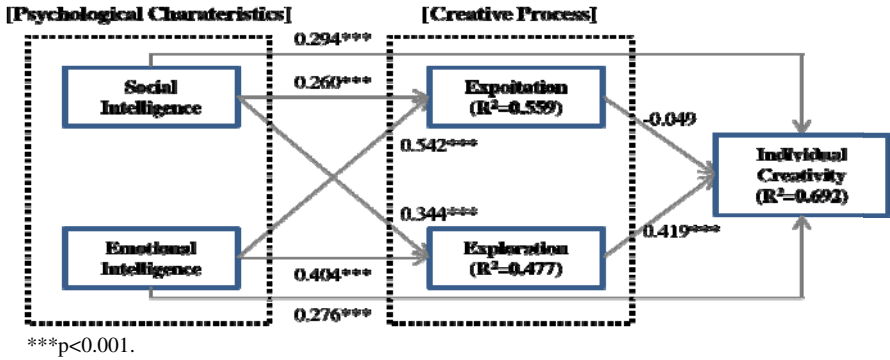


Fig. 1. Research Model Results

After the validity test of the hypotheses, all except H7 were statistically significant and therefore were accepted. The interpretation of the test results for each hypothesis is as follows. First, H1 and H2, which argue that social intelligence positively affects exploitation and exploration, respectively, were accepted at a 99.9% confidence level and the path coefficients were 0.260 and 0.344, respectively. This indicates that social intelligence positively affects exploitation and exploration in a work environment that requires creativity, especially in customer service jobs like the sample in this study. Second, H4 and H5, which posit that emotional intelligence has positive relationships with exploitation and exploration, respectively, were accepted at a 99.9% confidence level, with path coefficients of 0.542 and 0.404, respectively. This means that, in the relationship between exploration, exploitation, and social intelligence, individuals with higher emotional intelligence are more capable of exploitation and exploration. Third, H7 was rejected, while H8, which states that exploration positively affects individual creativity, was accepted at a 99.9% confidence level and the path coefficient was 0.419. This suggests that simply using existing knowledge does not directly affect individual creativity. Only when the activity is done to seek and explore new knowledge will individual creativity be directly affected. Lastly, H3 and H6 were accepted at a 99.9% confidence level and the path coefficients were 0.294 and 0.276, respectively. In other words, individuals with high social and emotional intelligence, which are personal psychological characteristics, utilize existing knowledge and explore new knowledge relatively well. However, at the same time, we could not find a directly significant effect on individual creativity.

5 Discussion and Conclusion

In this research, we suggested a new research model of individual creativity from the two perspectives of the creative process and personal psychological characteristics. An empirical analysis was conducted, with the sample including staff working in the IT industry, an area in which individual creative activities are vigorously undertaken. The results show that the personal psychological characteristics, measured by social

intelligence and emotional intelligence, have significant effects on the creative process, which is composed of exploration and exploitation activities. Exploration, which does not utilize existing knowledge, was found to have a significant effect on individual creativity.

This could be treated as a very meaningful result in the sense that the creative process and individual creativity are strengthened as social intelligence and emotional intelligence are strengthened. In other words, from the perspective of an organization, carrying out personal education and creating a suitable atmosphere for each individual may enable an increase in the creativity of each member. Below is a discussion on the direction that creative process, social intelligence, and emotional intelligence must have.

First, higher social intelligence and personal psychological characteristics are associated with stronger individual creativity. A common characteristic of today's successful people is that they have high social intelligence. Developing one's social intelligence in order to read the emotions of other people in the social relations network can be helpful, and may also increase creativity, which is one of the most important ability factors. Second, higher emotional intelligence is associated with improved creative processes and individual creativity. This leads us to deduce that the ability to motivate or encourage employees to do their best is related to individual creativity. Therefore, for excellent job performance, an individual should find a method to improve his individual creativity by understanding and controlling his own and co-workers' emotions sufficiently by cultivating the ability to use such emotions as well as by giving self-motivation by setting goals. Third, exploration has a more direct effect on individual creativity than exploitation when viewed from the perspective of the creative process. Therefore, from the organization's point of view, helping utilize existing knowledge and making and setting up the environment in which individuals can engage in exploration are very important.

In this research we investigated how personal psychological characteristics and the creative process should be used to increase individual creativity for both individuals and businesses. The results are based on an empirical study in a business environment in which creativity is very important. As with any study, we must consider the limitations and future research issues. First, this research cannot be applied to employees in all industries because only employees in the IT industry were included in the present study. Therefore, in future research, we can expand the study to explore if there are differences in the relationship between personal psychological characteristics and creativity depending on the industry. A second limitation of this research is that social intelligence and emotional intelligence are the only factors considered to affect individual creativity. Therefore, a more systematic explanation could be given if other personal psychological characteristics are considered in future research.

Acknowledgment. This research was supported by WCU(World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0).

References

1. Audia, P.G., Goncalo, J.A.: Past success and creativity over time: A study of inventors in the hard disk drive industry. *Management Science* 52(1), 1–15 (2007)
2. Barchard, K.A.: Emotional and social Intelligence: Examining its place in the nomological network. Unpublished Doctoral Dissertation, University of British Columbia (2001)
3. Chin, W.W.: The Partial Least Squares Approach to Structural Equation Modeling. In: Marcoulides, G.A. (ed.) *Modern Methods for Business Research* 1998, pp. 295–336. Lawrence Erlbaum Associates, NJ (1998)
4. Ciarrochi, J., Forgas, J., Mayer, J. (eds.): *Emotional intelligence in everyday life: A Scientific Inquiry*, vol. 1, pp. 157–188. Psychology Press/Taylor & Francis, Philadelphia, PA (2001)
5. Davies, M., Stankov, L., Roberts, R.D.: Emotional intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology* 75, 989–1015 (1998)
6. Fabio, S.: Leadership in Education: Effective UK College Principals. *Nonprofit Management and Leadership* 14(2), 171–189 (2003)
7. Gilson, L.L., Shalley, C.E.: What leaders need to know: A review of social and contextual factors that can foster or hinder creativity. *The Leadership Quarterly* 15, 33–53 (2004)
8. Goleman, D.: *Emotional Intelligence*. Bantam Books, NY (1995)
9. Goleman, D.: *Social Intelligence: The New Science of Human Relationships*. Bantam Books, NY (2006)
10. Guilford, J.P.: Creativity. *American Psychologist* 5, 444–454 (1950)
11. Guilford, J.P.: *The nature of human intelligence*. McGraw-Hill, NY (1967)
12. He, Z.L., Wong, P.K.: Exploration vs Exploitation: An Empirical Test of the Ambidexterity Hypothesis. *Organization Science* 15, 81–94 (2004)
13. Isen, A.M., Daubman, K.A., Nowicki, G.P.: Positive Affect Facilitates Creative Problem Solving. *Journal of Personality and Social Psychological* 52, 122–131 (1987)
14. Levinthal, D.A., March, J.G.: The myopia of learning. *Strategic Management Journal* 14, 95–112 (1993)
15. March, J.G.: Exploration and exploitation in organizational learning. *Organization Science* 2, 71–87 (1991)
16. Nerkar, A.: Old is good? The value of temporal exploration in the creation of new knowledge. *Management Science* 49, 211–229 (2003)
17. Salovey, P., Mayer, J.D.: Emotional intelligence. *Imagination, Cognition and Personality* 9(3), 185–211 (1990)
18. Sosik, J.J., Kahai, S.S., Avolio, B.J.: Leadership style, anonymity, and creativity in group decision support systems: The mediating role of optimal flow. *Journal of Creative Behavior* 33, 227–256 (1999)
19. Thorndike, E.L.: Intelligence and Its Uses. *Harpers Magazine* 140, 227–235 (1920)
20. Wong, C.M.T., Day, J.D., Maxwell, S.E., Meara, N.M.: A Multitrait-Multimethod Study of Academic and Social Intelligence in College Students. *Journal of Educational Psychology* 87, 117–133 (1995)
21. Wong, C.S., Law, K.S.: The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *The Leadership Quarterly* 13, 243–274 (2002)

An Management of Digital Contents on Science and Technology

Won-Goo Lee, Min-Ho Lee, Kwang-Young Kim, Sung-Ho Shin,
Hwa-Mook Yoon, and Won-Kyung Sung

KISTI, 245 Daehangno, Yuseong-gu, Daejeon, Korea
{wglee, kokeman, kykim, maximus, hmyoon, wksung}@kisti.re.kr

Abstract. It needs a management system that managers manage more conveniently, faster, more systematically and Users use faster, more various, more reliable, more discriminatory, higher value-added services on science & technology informations. Thereupon, We design and implement DiCMS(Digital Contents Management System) standardized and unified in a variety of literature data.

Keywords: digital contents, information management, contents management process.

1 Background

With We have managed information Management Systems such as NDSL(National Discovery for Science Leaders), OCEAN(Online Collaborator for sociEty & Association Network of KISTI), WISECAT(WISE CATalog), KESLI(Korean Electronic Site License Initiative) and KSCI(Korea Science Citation Index Service).

But, we have expected systematical and standardized digital contents management system considering semantic data as well as literature data(papers, articles, patents, reports, standards, trends) managed by KISTI to maximize the efficiency of the operation and management of data.

We also needed to manage and archive all digital contents including scientific data & semantic data as systematical and standard form. And we are constructing digital contents management system(DCMS) so that managers can manage efficiently, fast, systematically digital contents and users use fast, various, reliable, different, high-value services. To do this, we constructed the process and system to manage papers.

2 Intergration and Standardization

DCMS is management system on the variety of contents. We need the standardization and unification of systematic information management, the integration of a distributed database, and the standardization of items, words, domain, meta and data. We discuss the standardization of database and management process to manage digital content at this section and the creation and management of semantic data for smart service.

2.1 Integration

We disused a duplicated database itself and removed tables and fields unused on an existing database(including 400 tables, unified a various character sets(e.g, KSC5601 etc.,) to unicode format, standardized control number system by bestowing normally control number and processing systematically special character.

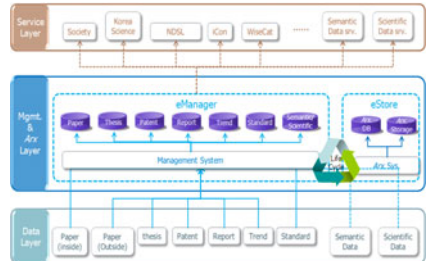
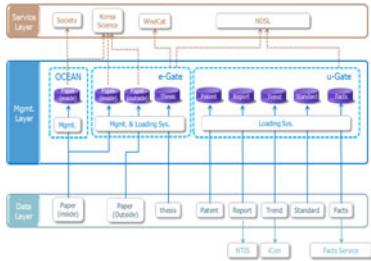


Fig. 1. An existing Management/Database Structure **Fig. 2.** Our Management/Database Structure

2.2 The Standardization of Data Format

We define the standard of data format with XML for all contents(papers, articles, patents, reports, trends, standards) in DCMS(Digital Contents Management System. We determine the method keeping pace with both database(control item and our XML standard considering the efficiency of our system, but an existing method with either database or a normal file(XML, MARC, plaintext format etc..).

We describe only the standard of data format for papers in this section. Paper(paper content) consist of journal information and article information. We select MARC XML format as our data format, because the main considerations for a management method of the metadata on journal information is to transfer lossless an existing MARC data into our XML standard. The MARC XML is the structure that it can be managed to XML based metadata and transferred lossless an existing MARC dat as the structure storing MARC data to XML format of the original form.

At the metadata of article information, We determined the standard of PMC(PubMed Central) XML which support XML schema module with a variety of elements and attributes, can be describe as full-text form and publish freely digital form such as e-book format and combine among the schema of completely different form as our XML standard. The reason that we determine PMC XML standard is of it to be easy to spread fast out and to select by associations and vendors, to be stabilized and have great authority as public domain, also to be easy to maintenance and management. We expand (<place>, publication location), (<addr>, publisher address or author/affiliation adress), (<frequency>, publication cycle), (<free-citation>, reference information) in PMC XML schema.

2.3 The Standardization of Process

We standardized and unified an existing content management process as well as the integration and standardization of database schema and data format. To this, we grasped an existing content management process, obtained new requirements from customers(content managers) through the analysis of the present state of their tasks and their needs.

First, the management process for the domestic papers consist of multi-step(acquisition - ordering - delivering - Checking - loading - managing - indexing step). Our system input or output in a temporary database from acquisition step to checking step, then it input or output in a regular database from loading step to indexing step. But, The management process for foreign papers is different from that for domestic papers. That is, the management process for the foreign papers consist of four steps(acquisition, loading, managing, indexing step). Likewise, we defined the process of other contents(thesis, patent, report, trend, standard content).

We construct DiCMS(Digital Contents Management System) based management process and system design, also implement our system followed web standards and based structs framework and java Script framework.

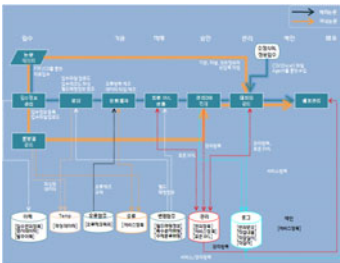


Fig. 3. Our Management Process on Papers

번호	제목	저자	등록일	상태	비고
1	국내논문	김민준	2010-01-15	완료	
2	해외논문	이영희	2010-02-20	완료	
3	국내논문	박지현	2010-03-10	완료	
4	국내논문	정민준	2010-04-05	완료	
5	국내논문	최민준	2010-05-01	완료	
6	국내논문	한민준	2010-06-01	완료	
7	국내논문	김민준	2010-07-01	완료	
8	국내논문	이영희	2010-08-01	완료	
9	국내논문	박지현	2010-09-01	완료	
10	국내논문	정민준	2010-10-01	완료	

Fig. 4. Our Process UI in DiCMS

3 Conclusions

We design and implement DiCMS(Digital Contents Management System) standardized and unified in a variety of literature data. To do this, we define the standard of management elements and XML standard as well as database to improve the efficiency of management. That is, we define information management, unify management system, integrate database distributed. At future, we will construct more systematical data management platform on semantic data, also research on the scientific data management and define its process standard.

References

1. Wongoo, L., Hwamook, Y., WonKyung, S.: An Efficient Management System on Digital Contents. In: ICC 2010, pp. 211–212 (2010)
2. Pyung, K.: OntoURI: Ontology Instance Management System based on URI. In: ICC 2008 (2008)

A Categorization and Solution of Metadata Conversion Problem on Research Paper

Min-Ho Lee¹, Hwa-Mook Yoon¹, Won-Kyung Sung¹, and Jae-Cheol Ryou²

¹ KISTI, 245, Daehakro, Yuseong-gu,
305806 Daejeon, Korea

² Chungnam National University, 99, Daehakro, Yuseong-gu,
305764 Daejeon, Korea

{cokeman, hmyoon, wksung}@kisti.re.kr, jcryou@cnu.ac.kr

Abstract. To conduct High level service of Science and Technology Information such as a deep forecast and analysis, and a knowledge discovery, we have to integratively manage and deal with various types of the information. In order to do that, making of integrated schema, schema mapping for data interoperability, data conversion and continuous managing should be preceded. In this paper, we will categorize the problems when metadata of research papers are converted, and describe a converting method for research papers specially.

Keywords: Ontology, Metadata, Conversion.

1 Introduction

Over the past service which simply provided information searching or distributing, in these days, information distributing companies are struggling to provide higher level service such as a deep forecast and analysis, and a knowledge discovery using various information processing technologies. Science and technology information can be roughly divided into scientific raw data such as experimental data, memo, research note and scientific literatures, which are research papers, patents and reports based on the raw data. To provide high level information service, these various kind of scientific information should be integratively managed and dealt. We, first of all, have to make integrated schema, connect between schemas for interoperation, convert data and continuously manage the integrated schema in order to that. In this paper, we will categorize the problems when metadata of research papers are converted, and describe a metadata converting method for research papers specially.

2 Characteristics of Elements in Research Paper Metadata for Converting

Metadata of research papers are respectively different from each publisher. They make new metadata schema for their own needs, but they mostly expand or transform standards metadata such as Dublin Core, MARC and ISO12083 [1]. Formats for

presenting the contents are mainly SGML or XML, but there are also plain texts (tagged text for the rest of this paper) which are defined by for themselves or HTML. This paper will not mention about structural or semantic problems when schema mapping [2]. The followings are varieties of content problems when content mapping for only research paper.

- 1) Code Conversion: There are many elements being representing by code. For examples: paper publish country element using country code, written language element using language code, topic element using classification code, publish period using season code, character code set and so forth.
We have to change codes for integration.
e.g.) Country code conversion
- China (in IEEE country code set) to CHN (in ISO3166 country code) [3]
- 2) Pattern Conversion: These types of elements are not be managed with code set but have regular patterns. For examples: Document identification, surname and given name of author, volume and issue, publish date, pages and so forth.
e.g.) Volume and Issue conversion
- 13th edition, V.1 to Vol. 1 No. 13
- 3) Other Conversion: This case is irregular content conversion that cannot be defined as code sets or patterns, or conversion between elements with slightly different semantics. For examples: changing page scope to page count, throwing out meaningless part of the content.
e.g.) Page scope conversion
- i-v (roman expression of page number) -> 5 (page count only)

3 Metadata Conversion Method by Conversion Type

This paper does not mention about user interface for schema mapping. The following methods by conversion type which is mentioned about section can convert any metadata without any system modification even if new schema enters into the system.

- 1) Code Conversion: After pre-registration of all of the code sets, letting user designate the source and target code set can easily convert this type of elements.
- 2) Pattern Conversion: String pattern can be applied to this type of elements. For example, if there is a book that has '1998' as volume number and issue '13' as issue number which is represented 'v.1998 no.13' as the element, the system can extract '@@@@' and '##' from whole string of 'v.@@@@ no.##' as a pattern. If user set string 'Vol.@@@@ no.##' as a output pattern, then the system can add the strings 'Vol.' and 'no.' to the extracted numbers '@@@@' and '##' and convert them.

3) Other Conversion:

In the case of other conversion, the system can allow users to register APIs in script language which have one input parameter and one output parameter.

For example, for presenting a specific paper in a journal, start page and end page of the paper in a source schema is expressed as form *i-v*. A target schema, in other hands, uses page counts only. This conversion is difficult with only code sets or patterns. We can convert with script language instead as following.

```
Read ("@@-@@", $1, $2);
$start = atoi($1); $end = atoi($2);
$pg = $end - $start + 1;
print $pg to page element of destination database;
```

Users write simple scripts for each element conversion if it is in other conversion case. The system can read the script conversion in the metadata conversion process, call script engine and do conversion.

4 Conclusion

We found some problems in the terms of content conversion and explain conversion method in this paper when metadata of research paper with various schemas were converted to unified metadata for integrated management of science and technology information. In addition to content conversion, we are planning to survey and categorize problems of semantic or structural collision when schema mapping and data converting. It is need to develop tool that is enable to easily map schemas and convert data without system modification. The method which automatically maps schemas will have to be studied.

References

1. Rosenblum, B., Golfman, I.: E-Journal Archive DTD Feasibility Study. Inera Inc. (2002)
2. Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. The VLDB Journal 10(4), 334–350
3. Country code, http://en.wikipedia.org/wiki/Country_code

Multi-words Terminology Recognition Using Web Search

Sa-Kwang Song, Yun-Soo Choi, Hong-Woo Chun, Chang-Hoo Jeong,
Sung-Pil Choi, and Won-Kyung Sung

Korea Institute of Science and Technology Information, Daejeon, Korea
{esmallj, armian, hw.chun, chjeong, spchoi, wksung}@kisti.re.kr

Abstract. Terminology recognition system which is a fundamental research for Technology Opportunity Discovery (TOD) has been intensively studied in limited range of domains, especially in bio-medical domain. We propose a domain independent terminology recognition system based on machine learning method using dictionary, syntactic features, and Web search results, since the previous works revealed limitation on applying their approaches to general domain because their resources were domain specific. We achieved F-score 80.4 and 6.4% improvement after comparing the proposed approach with the related approach, C-value, which has been widely used and is based on local domain frequencies. In the second experiment with various combinations of unithood features, the method combined with NGD(Normalized Google Distance) showed the best performance of 81.5 on F-score. We applied two machine learning methods such as Logistic regression and SVMs, and got the best score at SVMs method.

Keywords: Terminology Recognition, Text Mining, Machine Learning, Information Extraction, Technology Opportunity Discovery.

1 Introduction

Technology Opportunity Discovery (TOD) is an activity detecting technical emergences by analyzing meaningful information in literatures such as papers and patents to support strategic decision-making effectively, with the help of many technologies like Natural Language Processing, Text Mining, Semantic Web, etc. Recently, many researchers have been worked for supporting people who have decision making authority to make strategic technology plans [1, 13]. Usually, their tasks or systems import text mining methodologies to analyze tacit information inside company or on the Internet. Especially, technology entity recognition or technological terminology recognition is one of the most important research topics for the Technology Opportunity Discovery. In other words, the recognition of terminology representing *Technology* in documents is the fundamental part for further analysis such as relation analysis between *Technologies*, semantic web based analysis, and so on. In this paper, we propose a methodology for identifying *Technological* terminology in patents, which utilizes statistics in Web search results.

Many researches on terminology recognition in literatures like papers, patents, or reports have been introduced especially in biomedical domain [6,8,9,10]. Most of

them utilized *local* statistics obtained from the target corpus and some of them have proposed machine learning based approaches in addition to the *local* statistics based method [12]. *MI*(Mutual Information) [2], *LLR*(Log Likelihood-Ratio) [4] and *DC*(Dice Coefficient) [5] are the statistical terminology identification methods especially for calculating *unithood* which is an adhesion measure between two words. *C-value* [9] is one of the most popular methods, which utilizes the list of candidate terminologies in a corpus.

$$C\text{-value}(a) = \begin{cases} \log_2|a| \cdot f(a) & a \text{ is not nested,} \\ \log_2|a|(f(a) - \frac{1}{P(T_a)} \sum_{b \in T_a} f(b)) & \text{otherwise} \end{cases}$$

where a is the candidate term string, $f(\cdot)$ is its frequency of occurrence in the corpus, T_a is the set of extracted candidate terms that contain a , $P(T_a)$ is the number of these candidate terms. It is obvious that *C-value* is also a measure based on the *local* statistics of candidate terms.

The *local* statistics based terminology extraction might be domain dependent and is hard to respond the rapid change in technology terms such as neologism, time-worn words, or dead words. Therefore, we propose a machine learning based approach that utilizes *global* statistics obtained from web search engine. The *global* statistics in contrast to the *local* is domain independent and doesn't suffer from data sparseness because of the incomparable huge amount of data compared to domain specific corpus. In addition, the web statistics evolves as time goes by, so that it helps to capture the concurrent trends or popularity of technology.

To capture the *global* statistics, we utilized Normalized Google Distance [11]. It is a useful *unithood* method for terminology recognition using web statistics.

$$NGD(x, y) = \frac{\max\{\log f(x), \log f(y)\} - \log f(x, y)}{\log M - \min\{\log f(x), \log f(y)\}} \quad (1)$$

where $f(\cdot)$ is document frequency in Google search results, M is the total number of documents.

2 Methodology

Automatic recognition of terminology is divided into three main steps in general: terminology candidate extraction in a sentence, *unithood* (adhesion of all words in a candidate term) calculation for ranking candidates, and *termhood* calculation to decide whether a candidate term is a terminology or not according to various clues or information. For extracting terminology candidates, we used pattern based method which identifies noun phrases in a sentence using regular expression. Table 1 shows the patterns for noun phrase, noun phrase with preposition, and noun phrase with conjunction.

Table 1. Patterns in regular expression for candidate extraction

Pattern type	Patterns in regular expression	Pattern group
0	(JJ+ VB[GIN]+)*(NN[SIP]?+)+	Noun phrase ex) gene expression
1	((JJ+) (VB[GIN]))*(NN[SIP]?+)+(IN)(PRJ\$ DT)?(JJ+ VB[GIN])* (NN[SIP]?+)+ (Adj NN)? NN+(IN)	Noun phrase with preposition ex) expression of polypeptide
2	((JJ+) (VB[GIN]+))+(CC)((JJ+) (VB[GIN]+))* (NN[SIP]?+)+	Noun phrase with conjunction
3	(NN[SIP]?+)+(CC)((JJ+) (VB[GIN]+))* (NN[SIP]?+)+	ex) B and T cell

After collecting candidates from sentences or documents according to the patterns in table 1, we calculated the *unithood* of each candidate using MI, LLR, DC, NGD. And then additional features for each candidate term in addition to the *unithood* are gathered and applied to decide whether it is a terminology or not using machine learning approach. After all, the *termhood* score for a candidate term is calculated and then the term is decided whether it is terminology or not through this machine learning method using various features including *unithood* measures.

We applied logistic regression [7] and SVMs (libSVM¹) [3] methods using Weka toolkit² and the additional features extracted for terminology recognition are as follows.

① Syntactic Features (S) <See table 1>

- Pattern type: pattern style applied on extracting candidates
- Pattern group: noun phrase, noun phrase with preposition, and noun phrase with conjunction
- # of words in candidate

② Dictionary (D)

- Whether the candidate is in terminology dictionary
- Whether web search result with a candidate as a query contains Wikipedia page

③ Web features (W)

- Web frequency (minimum, maximum, and average) from search results of individual terms in a candidate
- Whether the search results of individual terms in the candidate contains Wikipedia page
- # of individual terms that contains any Wikipedia page in the search result

④ C-value (C)

⑤ Normalized Google Distance (F1)

¹ libSVM, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

² Weka, <http://www.cs.waikato.ac.nz/ml/weka/>

- ⑥ Mutual Information (F2)
- ⑦ Log-likelihood Ratio (F3)
- ⑧ Dice Coefficient (F4)

③,⑤,⑥,⑦, and ⑧ are based on web search results that reflects the concurrent web statistics. For efficiency, we cached the statistics obtained from web search results for later use. *Unithood* (⑤,⑥,⑦, and ⑧) from the web results is calculated like this. Suppose that we found a chunk consisting of four words, 'A B C D'. It could have three candidates like 'A B C D', 'B C D', and 'C D' when D is a head word. Each of them then is sent to a web search (Google, Bing, etc). For example, 'B C D' is a candidate word to calculate the *unithood*, and 'B', 'C D', and 'B C D' are the very queries to be sent to the search engine. Then we could get $f('B')$, $f('C D')$, and $f('B C D')$ for the NGD formula in (1). In this way, the scores for ⑥,⑦, and ⑧ could be obtained.

3 Experiments

The data set for experiment consists of 300 papers and 100 American patents. The average sizes of the paper and the patent are 1.95KB and 6.28KB respectively. It has constructed by two professional terminologists and 7,118 terminologists are tagged eventually. Using the features introduced in section 2, we trained and tested the data by two machine learning approaches such as logistic regression analysis and SVMs. For kernel function for SVMs, we use the radial basis function since it has outperformed the other three kernel functions provided by the libSVM toolkit. The parameter optimization for the SVMs, grid.py script included in the toolkit is applied.

As the first experiment, we compared four features in addition to dictionary feature. The syntactic feature or web feature based experiments (DS, DW, DWS) are outperformed C-value method (C, or DC). To our surprise, the C-value addition to the DWS (the best performed setting) shows degradation in SVMs method.

Table 2. Comparison of the features, C(C-value), Syntactic features(S), Web features(W), and Dictionary(D). DWS means the three features (D,W,and S) are used for machine learning approach.

		<i>D</i>	<i>C</i>	<i>DC</i>	<i>DS</i>	<i>DW</i>	<i>DWS</i>	<i>DWSC</i>
Logistic Reg.	Precision	57.6%	72.1%	75.4%	78.2%	78.9%	79.1%	79.2%
	Recall	75.9%	75.9%	76.2%	78.2%	79.7%	79.8%	79.9%
	F-measure	65.5%	74.0%	75.8%	78.2%	79.3%	79.4%	79.5%
SVMs	Precision	57.6%	69.0%	74.8%	78.7%	79.8%	80.6%	77.9%
	Recall	75.9%	75.8%	76.7%	77.8%	79.8%	80.2%	79.5%
	F-measure	65.5%	72.2%	75.7%	78.2%	79.8%	80.4%	78.6%

Based on the first experiment, we compared the effect of four *unithood* methods, and got the best when NGD (F1) is added to the best setting in the first experiment. In this experiment, we also added the C-value feature to the best setting (DWSF1), but again it shows decrease in performance. This experiments show that the web statistics based approach could replace the local statistics based approach such as C-value.

Table 3. Comparison of the four *unithood* features, NGD(F1), MI(F2), LLR(F3), and DC(F4)

		<i>DWSCF1</i>	<i>DWSF1</i>	<i>DWSF2</i>	<i>DWSF3</i>	<i>DWSF4</i>
Logistic Reg.	Precision	80.4%	80.6%	79.6%	79.1%	79.8%
	Recall	80.6%	80.7%	80.4%	79.8%	79.9%
	F-measure	80.4%	80.6%	80.0%	79.4%	79.8%
SVMs	Precision	79.4%	81.9%	81.3%	79.3%	68.9%
	Recall	80.6%	81.1%	81.1%	79.8%	75.4%
	F-measure	80.0%	81.5%	81.2%	79.5%	72.0%

4 Conclusion and Future Work

We proposed web statistics based terminology recognition approach which does not depend on local statistics. This web based method could be independent to specific domain and could be appropriate for adopting the system to the continuously changing technology trends. We utilized two machine learning methods, logistic regression and SVMs, and got the best performance from SVMs method with NGD (Normalized Google Distance), syntactic features, web features, and dictionary. The best F-score was 81.5 which is quite promising. For future work, we are going to utilize the classification or cluster information such as Wikipedia category.

References

1. Yoon, B.: On the development of a technology intelligence tool for identifying technology opportunity. *Expert Systems with Applications* 35, 124–135 (2008)
2. Church, K., Hanks, P.: Word association norms, mutual information, and lexicography. *Computational Linguistics* 16(1), 22–29 (1990)
3. Cortes, C., Vapnik, V.: Support-Vector Networks. *Machine Learning* 20(3), 273–297 (1995)
4. Dunning, T.: Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1), 61–74 (1993)
5. Smadja, F., McKeown, K.R., Hatzivassiloglou, V.: Translating collocations for bilingual lexicons: A statistical approach. *Computational Linguistics* 22(1), 1–38 (1996)
6. Wermter, J., Hahn, U.: Paradigmatic Modifiability Statistics for the Extraction of Complex Multi-Word Terms. In: *HLT 2005 Proceedings of the Conference on Human Language Technology and Empirical Methods in NLP* (2005)
7. Hilbe, J.M.: *Logistic Regression Models*. Chapman & Hall/CRC Press (2009)
8. Justeson, J.S., Katz, S.M.: Technical terminology: some linguistic properties and an algorithm for identification in text. *Natural Language Engineering* 1(1), 9–27 (1995)
9. Frantzi, K., Ananiadou, S., Mima, H.: Automatic recognition of multi-word terms: the C-value/NC-value method. *International Journal on Digital Libraries* 3(2), 115–130 (2000)
10. Nakagawa, H., Mori, T.: Automatic term recognition based on statistics of compound nouns and their components. *Terminology* 9(2), 201–219 (2003)

11. Cilibrasi, R., Vitanyi, P.: The Google Similarity Distance. *IEEE Trans. Knowledge and Data Engineering* 19(3), 370–383 (2007)
12. Zeng, Q.T., Tse, T., et al.: Term identification methods for consumer health vocabulary development. *Journal of Medical Internet Research* 9(1) (2007)
13. Tseng, Y., Lin, C., Lin, Y.: Text mining techniques for patent analysis. *Information Processing and Management* 43(5), 1216–1247 (2007)

Disambiguating Author Names Using Automatic Relevance Feedback

In-Su Kang^{1,*} and Seung-Hoon Na²

¹ Computer Science and Engineering, Kyungsoong University, Pusan 608-736, South Korea

² Department of Computer Science, National University of Singapore

dbaisk@ks.ac.kr, nash@comp.nus.edu.sg

Abstract. Author disambiguation is to associate author-name instances with author identifiers corresponding to real-world authors. Bibliographic records, which normally contain author-name occurrences to be disambiguated, are not sufficient to provide author-resolving features. This study proposed a feedback-based method for author disambiguation to tackle the insufficiency of author representation. Our method is focused on enhancing author's topical representation by adding topically-related terms obtained from feedback documents. Experiments showed the positive effect of the method on author resolution by improving the performance.

Keywords: Author Disambiguation, Automatic Relevance Feedback, Web Evidence.

1 Introduction

Author disambiguation is to map author-name occurrences into real-world individuals. Its difficulty comes from many-to-many relationships between names and persons. It can improve the quality of academic search systems by enabling author-oriented grouping of bibliographic records rather than name-oriented.

In general, a bibliographic record in which author names to be disambiguated appear consists of co-author name(s), a paper title, a publication title, a publication year, etc. However, such information is not enough to provide person-identifying features to author disambiguation systems. Researchers have thus attempted to supplement this with information from either full-text archives or personal publication-list pages on the Web. These external sources however are not always or mostly available.

This study tries to attack the above problem by expanding the basic bibliographic record with its topically-related articles obtained from a large collection of bibliographic records for which there exist many freely-accessible archives or services such as DBLP and PubMed. For this, we apply the relevance feedback technique well-developed in information retrieval (IR) communities. Experiments show that our method effectively improves the performance of author disambiguation

* Corresponding author.

by 34.6%. From the viewpoint that author-identifying features can be classified into biographic ones (e.g., coworkers, emails and affiliations) and topical ones (e.g., academic areas and research topics), we handle the latter.

The remainder of this article is organized as follows. Section 2 reviews the related works. Section 3 describes a feedback-based method of enriching author's topical features. Section 4 reports the experimental results, and Section 5 provides the conclusion.

2 Related Work

This section reviews the earlier uses of influential author-identifying features to enhance author representation. For an extensive survey of author disambiguation, readers can refer to [1].

Several researchers have reported on the successful application of PPLW (personal publication-list pages on the Web) features to author resolution [2-4]. PPLW indicate the web page that contains a list of bibliographic records authored by a particular researcher. Instead of directly identifying and gathering PPLW pages from the Web, previous works have attempted to indirectly find the evidence from the general search engine that the same web page includes two bibliographic records of which author names are to be disambiguated.

The full-text of an article is another rich source for improving author representation, which holds biographic features such as coauthors, email addresses and affiliations, as well as topical features such as keywords, abstracts, references, etc. Song[5] exploited the first page of the article document to train Latent Dirichlet Allocation model and observed the best disambiguation performance 93.6% for English names. Kang[6] improved the performance of Korean author disambiguation up to 91.47% in F1 by employing a combination of co-author names and emails extracted from the article full-text.

Unfortunately, both PPLW and the full-text of articles have some limitations on their accessibility and availability. In other words, author disambiguation systems cannot guarantee that all or most authors of articles would maintain their publication pages on the Web. Besides, all or more full-text archives are not freely accessible. Unlike such hardly-obtainable resources, the collections of bibliographic records that are exploited by this study are relatively easily reachable.

3 Feedback-Based Enhancement of Author Representation

Author disambiguation is normally applied to a collection of same-name author occurrences. Its general procedure is as follows. First, every name instance is converted into its author-instance representation. Next, for each pair of author-instance representations, its similarity is computed. As a result, a graph is obtained where a node indicates a name instance, and the link between two nodes corresponds to the similarity between two author-instance representations. Then, a clustering technique is applied to the graph to produce a set of author clusters of name instances.

This study focuses on enriching the author representation to get more accurate similarity values. In the following, a feedback-based procedure of enhancing author representation is described.

Let's suppose that a bibliographic record consists of co-author names $a_1, \dots, a_i, \dots, a_n$, paper title t , and publication title p , and that name a_i should be disambiguated. Then, a representation A of author name instance a_i can be divided into biographic representation $A_B = \{a_1, \dots, a_n\} - \{a_i\}$ and topical one $A_T = \{t, p\}$.

The steps to enrich topical representation A_T through automatic relevance feedback are as follows. First, query Q is formulated from A_T . Next, Q is submitted to an information retrieval system which is assumed to have indexed a large collection C of bibliographic records. In response to Q , a set F of top k bibliographic records is retrieved from C . Finally, original topical representation A_T is enriched from F . There may be several techniques to extend A_T from F . A simple method is to gather all paper titles from F and to union them with A_T to produce enhanced topical representation A_T' . Another more common way is to select top m terms from the set of entire terms extracted from F using a term-scoring formula, and to combine them with A_T to get A_T' . The above feedback loop can be repeated many times using the previous A_T' as a query to generate the next A_T' .

4 Experiments

The proposed method is evaluated using PSU-CiteSeer-14 dataset. It is comprised of a total of 8,453 bibliographic records which were gathered for 14 English person names and manually disambiguated into 479 individual identifiers.

For every name instance, its author (instance) representation is generated in the form of a feature vector from the bibliographic record to which the name instance belongs. To calculate the similarity between two author representations A^i and A^j , the following equation, a variant of that used in [7], is employed.

$$Sim(A^i, A^j) = \delta \left(\left(\sum_{k=1}^K \delta(fsim_k(A_k^i, A_k^j) \geq \theta_k) \right) \geq \pi \right). \quad (1)$$

$fsim_k(\cdot, \cdot)$ is the feature similarity between k -th features of A^i and A^j . $\delta(p)$ is defined to return 1 if p is true, 0 otherwise. θ_k is a threshold value of k -th feature. This experiment assumes that the author representation has two feature types: biographic and topical ones which correspond respectively to co-author names and paper/publication titles. Biographic-feature similarity $fsim_B(\cdot, \cdot)$ ¹ is defined to be the count of shared co-author name(s) between A^i and A^j . Topical-feature similarity $fsim_T(\cdot, \cdot)$ is the cosine similarity between paper/publication titles of A^i and A^j . For this, titles are decomposed into lowercased stemmed terms after stop-word removal, and each term is weighted by logarithmized tf-idf.

Google scholar was used as a large collection of bibliographic records to enhance author's topical representation. For each of 8,453 records, its paper title was

¹ $fsim_B(\cdot, \cdot)$ and later $fsim_T(\cdot, \cdot)$ can be understood as $fsim_1(\cdot, \cdot)$ and $fsim_2(\cdot, \cdot)$ in the notation of Formula (1).

submitted as a query to Google scholar to retrieve top 10 relevant articles. To get more relevant topical terms from feedback documents, two expansion methods described in Section 3 were attempted.

Following the experimental design of Pereira et al.[4] that reports the state-of-the-art author-resolving performance, our data set is randomly divided into train and test sets in halves. The train set is used to tune the parameters θ_k and π , and the test set to get recall, precision, and F1 figures. The above procedure is iterated ten times and their averaged performances are reported.

Table 1 shows evaluation results. Earlier experiments on author disambiguation mostly use HAC as a baseline method. However, BFC, a biggest-first clustering method recently proposed in [7], showed the better effectiveness than HAC (Hierarchical Agglomerative Clustering) based on single-linkage. So, all our later experiments employed BFC. Given a graph of entities to be clustered, BFC basically iterates the routine that removes a sub-graph g of the largest-degree node and its adjacent nodes from the input graph and registers g as a new cluster or merges g into one of existing clusters until the input graph becomes a null graph. $\text{FBr}(\mathbf{x}, \mathbf{y})$ indicates the use of automatic feedback which expands original paper/publication titles with top \mathbf{y} terms from top \mathbf{x} titles retrieved.

Table 1. Performance of author disambiguation ($\pi=\theta_b=1$)

Method	Recall	Precision	F1	Parameters
HAC (baseline)	0.4583	0.1348	0.2083	$\theta_l=0.61$
BFC (baseline)	0.2842	0.5028	0.3631	$\theta_l=0.3$
BFC+FBr(10,all)	0.3986	0.5231	0.4524	$\theta_l=0.32$
BFC+FBr(5,all)	0.3460	0.5358	0.4205	$\theta_l=0.32$
BFC+FBr(10,30)	0.2914	0.5411	0.3788	$\theta_l=0.32$
BFC+FBr(10,all)+FBc(*,all)	0.4112	0.6024	0.4887	$\theta_l=0.34$

The use of $\text{FBr}(10,all)$ led to a significant improvement over the case without feedback by 24.6%. This means that author's topical representation can be effectively enriched from highly relevant titles of other articles. The fact that $\text{FBr}(5,all)$ underperformed $\text{FBr}(10,all)$, however, implies that the effect of adding as many related terms to author representation overshadows that of eliminating noisy terms from possibly irrelevant titles. A similar result was clearly confirmed from the result of $\text{FBr}(10,30)$, the expansion using top-scoring 30 terms from top 10 article titles, yielding a small increase in precision but a large decrease in recall.

To score the term t from top titles in $\text{FBr}(10,30)$, we exploited the well-known relevance weight [8] $rw(t)$ shown in the below, where in IR terminology R is the number of relevant documents of the given query, N the number of total documents, $r(t)$ the document frequency of term t among relevant documents, and $n(t)$ the document frequency of term t . To approximate the values of these parameters that cannot be directly obtained from Google scholar, a DBLP dump² was used.

² <http://www.informatik.uni-trier.de/~ley/db/>

$$rw(t) = \log \frac{(r(t)+0.5)(N-n(t)-R+r(t)+0.5)}{(n(t)-r(t)+0.5)(R-r(t)+0.5)}. \quad (2)$$

From the lesson of earlier experiments, we have attempted to further enlarge the set of relevant article titles rather than to discard presumably erroneous expansion terms. For this, it was assumed that an article is relevant to the later-published article citing the former. Then, for each article in our test data, a collection of other articles citing the former article was gleaned from Google scholar. $FBC(*,all)$ in Table 1 indicates the expansion of *all* terms extracted from titles of such citing-articles (* means zero or more citing-articles). Expectedly, the addition of $FBC(*,all)$ showed better effectiveness, increasing both recall and precision simultaneously. This observation not only reconfirms the usefulness of our feedback-based technique applied to author disambiguation, but also shows the possibility that the proposed method could yield further improvements through the use of feedback documents with different characteristics. Overall, the automatic feedback from retrieved documents and citing-articles contributed to enhance the performance by 34.6% compared to BFC baseline.

Table 2 compares our evaluation result with those of previous works. Kang[7] reports 49.42% in F1 using co-citation information extracted from the sets of citing-articles which were gathered from Google scholar in the same way as ours. Combining co-citation feature with ours performed better than Kang's[7], resulting in 53.52% in F1. It implies that co-citation and title expansion come from the same source but each of them brings its own positive effect on author disambiguation. The lower three methods in Table 2 were evaluated in [4] as clustering approaches to author disambiguation, where WAD is a kind of HAC. PPLW means that the related methods have exploited personal publication-list pages on the Web as the author-resolving information source. Considering that PPLW is limited in covering arbitrary articles, our proposal may be a promising solution to general-purpose author disambiguation.

Table 2. Comparison with previous works[4,7]

Method	Recall	Precision	F1
BFC+FBr(10,all)+FBC(*,all)+Co-citation	0.4574	0.6447	0.5352
BFC+Co-citation	0.4200	0.6003	0.4942
HAC+PPLW	n/a	n/a	0.46
k-way spectral clustering	n/a	n/a	0.36
WAD+PPLW	n/a	n/a	0.84

5 Conclusion

This study attacked author disambiguation by enriching the basic bibliographic record with its topically-related documents. For this, automatic relevance feedback techniques were applied. Experiments showed the positive effect of the method on author resolution by improving the performance.

In this study, expansion using top-scoring terms from feedback documents reduced the recall, while enhanced the precision. This may be due to not using low-scoring terms in enhancing the topical representation. Later we would apply query reformulation techniques developed in IR community, which can reflect the weights of feedback terms such as the relevance weight to produce the enriched topical representation.

In general, the feedback loop can be iterated multiple times, while this work employed a single feedback loop due to the time delay from running Google scholar-based feedback scheme. In the future, multiple feedback loops could be attempted to not only gather more feedback documents but also elaborate the author's topical representation.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2011-0012099).

References

1. Smalheiser, N.R., Torvik, V.I.: Author name disambiguation. *Annual Review of Information Science and Technology* 43, 287–313 (2009)
2. Aswani, N., Bontcheva, K., Cunningham, H.: Mining Information for Instance Unification. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L.M. (eds.) *ISWC 2006*. LNCS, vol. 4273, pp. 329–342. Springer, Heidelberg (2006)
3. Tan, Y., Kan, M., Lee, D.: Search engine driven author disambiguation. In: *Proceedings of ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pp. 314–315 (2006)
4. Pereira, D., Ribeiro-Neto, B., Ziviani, N., Laender, A.: Using Web information for author name disambiguation. In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pp. 49–58 (2009)
5. Song, Y., Huang, J., Councill, I., Li, J., Giles, C.L.: Efficient topic-based unsupervised name disambiguation. In: *Proceedings of the ACM IEEE Joint Conference on Digital Libraries (JCDL)*, pp. 18–23 (2007)
6. Kang, I.S.: Email Extraction and Utilization for Author Disambiguation. *The Journal of the Korea Contents Association* 8(6), 261–268 (2008) (in Korean)
7. Kang, I.S.: Disambiguation of Author Names Using Co-citation. *Journal of Information Management* 42(3), 167–186 (2011) (in Korean)
8. Robertson, S.E., Sparck Jones, K.: Relevance Weighting of Search Terms. *Journal of the American Society for Information Science* 27(3), 129–146 (1976)

OntoURIResolver: URI Resolution and Recommendation Service Using LOD

Taehong Kim^{1,2}, Pyung Kim¹, Seungwoo Lee¹, Hanmin Jung¹, and Won-Kyung Sung¹

¹ Knowledge Information Center,
Korea Institute of Science and Technology Information, Korea

² Practical Information Science,
University of Science & Technology, Korea
{kimtaehong, pyung, swlee, jhm, wksung}@kisti.re.kr

Abstract. LOD (Linked Open Data) provides an infrastructure to exchange multiples ontologies with standard formats. LOD recommends some rules about data publishing, URI (Uniform Resource Identifier) assignment, URI reusing and ontology accessing. To make a local ontology interoperable and sharable with LOD, we have to make linkages URIs between local ontology and LOD. In order to decide appropriate URI for specific entity, OntoURIResolver collects RDF (Resource Description Framework) triples of multiple ontologies by SPARQL, divides URIs into several groups by comparing RDF triples of URIs and recommends a canonical URI and entity name for each group using statistics of RDF triples. We experiment comparison of sameas.org and OntoURIResolver with top 10 ranked authors of DBLP. Users can find a specific URI for entity and make interconnections with LOD to maximize the effectiveness of ontology through this service.

Keywords: OntoURIResolver, URI Resolution, Ontology, Linked Open Data.

1 Introduction

Ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. In order to share, reuse and unify data set, we use ontology as a knowledge expression model. To maximize the effectiveness of ontology, the relationships between classes, properties and instances are required. W3C provides principles about building, publishing and accessing ontology through LOD project and many ontologies such as FOAF, DBpedia, DBLP are cross-linked and being utilized actively.

The entities that make up the ontology are represented by URIs and URI are used as an identifier to obtain related properties and conjunction point to connect different sources. Identified relationships between URIs can be expressed by using owl:sameAs, owl:differentFrom and owl:AllDifferent in OWL (Web Ontology Language), but it is not easy to find appropriate URI for specific entity. Because there are too many equivalent URIs in LOD. And we can find some errors in co-referent relationships that were published in LOD and it propagates incorrect results.

In this study, we propose OntoURIResolver service for efficient instance integration of your own URIs and URIs in LOD by providing these functions; 1) User can get URIs by URI or entity name, 2) Service provides identified groups of URIs and recommends canonical URI and entity name for each group. 3) User can edit results generated by service and modified results by users will be used for URIs resolution next time. Currently, we collect URIs from sameas.org and sindice.com and gather RDF triples from SPARQL endpoints using web services. The resolved types of our service are limited to author and article and we are expanding resolved types gradually.

2 Related Work

Most researches for entity identification have been carried out to disambiguate the authors of academic literatures by NLP (Natural language Processing) techniques [1,2,3]. Large evaluation dataset was constructed and information obtained through a web search was used to enhance the performance of author disambiguation [4]. But these researches have limitation on disambiguation type. And Scopus provides academic information service using author disambiguation [5].

Southampton Univ. develops and operates sameas.org service that provides co-referent information collected from LOD and other ontologies [6]. They gathered co-referent information by SPARQL endpoints and RDF dumps. It is very useful service, because user can search URI lists and co-referent information which are spared widely in the web through sameas.org. But, they do not create new URI group using RDF triples and the collected co-referent data contains incorrect information.

DBpedia provides identifying information between the internal and external entities by representing these properties such as dbpedia-owl:wikiPageRedirects, dbpedia-owl:wikiPageDisambiguates and owl:sameAs. But this information is constructed by manually and there is no service to control co-referent information automatically.

3 OntoURIResolver

OntoURIResolver is a URI resolving service that collects RDF triples in LOD, makes groups per entity and recommends canonical URI and entity name per each group in order to link our own ontology to LOD. OntoURIResolver uses OntoReasoner [7], which is a reasoning engine developed by KISTI, as triple store, Resolving process is executed in real-time because, RDF triples in LOD are frequently changed and the results depend on the status of network and external services. It collects RDF triples by using SPARQL endpoints and accessing URIs. And it uses triple store as a temporarily to store collected RDF triples, mapping table of classes and properties between multiple ontologies, pre-executed resolved results and edited information by users.

The process of OntoURIResolver is shown as figure 1 and figure 2. Figure 1 shows the overall process of OntoURIResolver. The process can be divided into five stages; 1) Comparison to cache: To resolve URIs, we have to crawl RDF triples by web services and compare properties of URIs. These tasks are time consuming process and

total time depends on network traffics and responding time of multiples web services. If there are no changes of URIs in LOD and pre-executed resolved result is existed then, we do not have to resolve URIs again. When a user input accepted, the system invokes a list of URIs to sameas.org or sindice.com with User's query, which is URI or entity name. To prevent duplicate process and reduce unnecessary time for resolving same entity, a system use cache memory and compare URIs to be resolved with cached results before processing URI resolution. 2) URI Resolution: The details of this process are explained at figure 3. 3) There are many ontologies and equivalent URIs in LOD. To link our instance to others, we have to make relationships between URIs assigned to same group. The names of person or institution have many variants in multiple documents. Sometimes we have select typical entity name and URI for building an ontology. For example, KISTI is expressed with 'KISTI', 'Korea Institute of Information and Science Technology' and sometimes KISTI is written in other language. The longest entity name is selected to canonical name and canonical URI is selected by considering the number of RDF triples and URI having canonical name. 4) Visualization: The resolved result can be divided into several groups such as figure 3. When user request the RDF triples belonging to URIs of specific group, OntoURIResolver crawls RDF triples by accessing URI in real-time. 5) Edit Results: User can edit the resolved result and modified result. And these will be used for resolving URIs at next time.

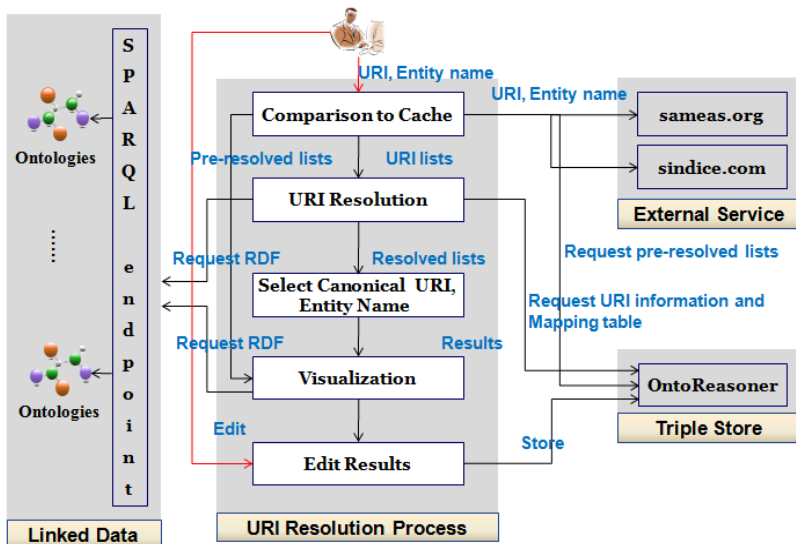


Fig. 1. Overall Process of OntoURIResolver

Figure 2 shows URI resolution process of figure 1 in detail. This process can be divided into six stages; 1) Collect RDFs of URIs: In this stage, URIs are classified into several groups according to status of URL, returned data types and status of URI. One is normal group containing URIs to be resolved and the other is abnormal group

containing deprecated URIs, dead links and URIs that formed in non-RDF data. After this stage, only normal group will be treated at next steps. 2) Make groups by entity type: we can separate some URIs using information type of URIs. To improve the accuracy of type judgment, we constructed the relationships between classes. 3) Make groups by entity name: URIs in normal group can be divided into several subgroups based on similarity of entity names that is calculated with string match method. If the threshold is high then some URIs remains separately. And if the threshold is low then this method makes over-clustering. We have to enhance the accuracy of this method with various heuristics. 4) Collect additional properties of URIs: In this stage, OntoURIResolver collects additional RDF triples to divide URIs having same type and same name. For resolving authors, we use some properties related to 'creator', 'author', 'articles created by someone', 'email', 'affiliation' and etc. To do this process, we constructs mapping table about relationships between properties of LOD. To expand resolving types, more information related to various types are required. RDF triples are obtained by SPARQL endpoints in real-time. 5) Compare multiples properties of URIs: RDF triples collected in previous stage is compared to each other. 6) Identify URIs: URIs are separated to some groups using previously calculated values.

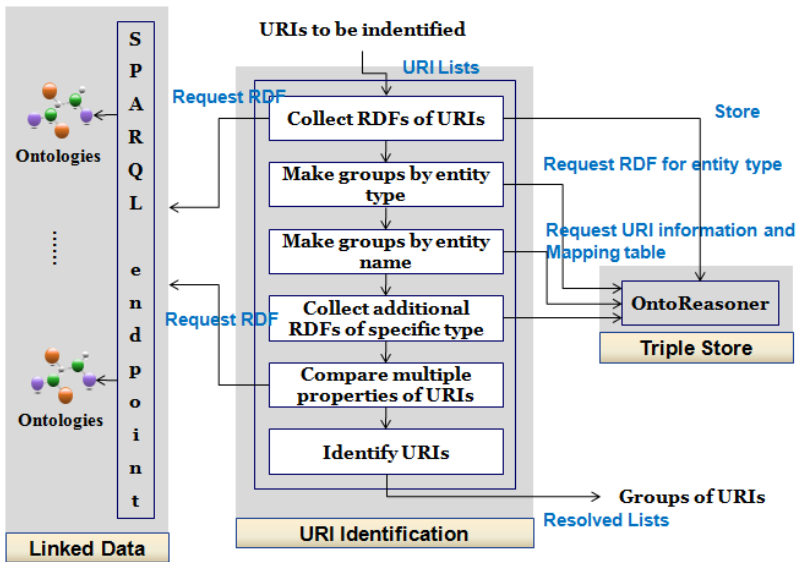


Fig. 2. Process of URI Resolution

Figure 3 is a snapshot of OntoURIResolver service. The service shows different results of OntoURIResolver and sameas.org. Actually, sameas.org does not try to merge URIs. They provide only co-referent information gathered by LOD. Of course, sameas.org provides very useful information about co-referent relationships between URIs. But OntoURIResolver classifies normal URIs and abnormal URIs and helps user to select URIs of specific entity. User can see RDF triples belonging to specific group by clicking button.

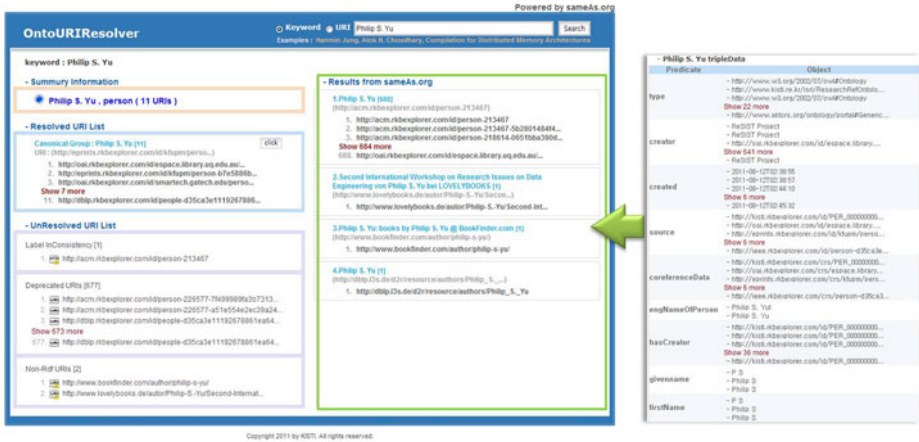


Fig. 3. User Interface of OntoURIResolver

4 Experiments and Evaluation

Our proposed method is experimented by using top 10 ranked authors of DBLP. Top ranked author means a person who has many publications in DBLP. Generally, there are many RDF triples and URIs related to top ranked authors. And then we compared our resolved results to sameas.org. The number of groups created by OntoURIResolver is smaller than sameas.org and the correctness of our results is checked manually.

Table 1. Comparison Table of sameas.org and OntoURIResolver

Author Name	sameas.org		OntoURIResolver		
	# of group	# of URI	# of group	# of valid URI	# of invalid URI
Philip S. Yu	4	690	1	11	689
Chin-Chen Chang	9	80	1	11	69
Elisa Bertino	8	532	1	15	517
Thomas S. Huang	10	28	1	5	23
Wen Gao	10	12	2	3	9
Wei Zhang	10	12	1	6	6
Edwin R. Hancock	3	570	2	9	561
Sudhakar M. Reddy	3	588	1	11	577
Ming Li	10	10	1	3	7
Li Zhang	10	10	2	6	4

5 Conclusions and Future Work

Ontology is very useful knowledge expression model for sharing, reusing and unification of multiple data sources. And LOD provides effective environment to interconnection among multiple ontologies. To make a local ontology interoperable

and sharable with LOD, we have to make linkages URIs between local ontology and LOD. In order to decide appropriate URI for specific entity, OntoURIResolver collects RDF (Resource Description Framework) triples of multiple ontologies by SPARQL, divides URIs into several groups by comparing RDF triples of URIs and recommends a canonical URI and entity name for each group using statistics of RDF triples. We experiment comparison of sameas.org and OntoURIResolver with top 10 ranked authors of DBLP. Users can find a specific URI for entity and make interconnections with LOD to maximize the effectiveness of ontology through this service. In this study, we proposed OntoURIResolver service for efficient instance integration of your own URIs and URIs in LOD. The resolved types of our service are limited to author and article and we are expanding resolved types gradually.

References

1. Gentile, A.L., Zhang, Z., Xia, L., Iria, J.: Semantic Relatedness Approach for Named Entity Disambiguation. *J. Communications in Computer and Information Science* 91(4), 137–148 (2011)
2. Pilz, A., Paaß, G.: Named entity resolution using automatically extracted semantic information. In: *Entity 2010 Workshop at LREC 2010, Malta* (2010)
3. Brill, E.: Pattern-Based Disambiguation for Natural Language Processing. In: *EMNLP/VLC-2000, HongKong* (2000)
4. Kang, I., Na, S., Lee, S., Jung, H., Kim, P., Sung, W., Lee, J.: On co-authorship for author disambiguation. *J. Information Processing & Management* 45(1), 84–97 (2009)
5. Elsevier B.V., "SciVerse ScopusSciVerse" (2011), <http://www.info.sciverse.com/scopus/>
6. Glaser, H., Jaffri, A., Millard, I.C.: Managing Co-reference on the Semantic Web. In: *WWW 2009 Workshop: Linked Data on the Web (LDOW 2009), Madrid* (2009)
7. Lee, S., Lee, M., Kim, P., Jung, H., Sung, W.: OntoFrame S3: Semantic Web-Based Academic Research Information Portal Service Empowered by STAR-WIN. In: Aroyo, L., Antoniou, G., Hyvönen, E., ten Teije, A., Stuckenschmidt, H., Cabral, L., Tudorache, T. (eds.) *ESWC 2010. LNCS, vol. 6089*, pp. 401–405. Springer, Heidelberg (2010)

A Comparison of Classifiers for Detecting Hedges

Sin-Jae Kang¹, In-Su Kang², and Seung-Hoon Na³

¹ School of Computer and Information Technology, Daegu University
Gyeongsan, Gyeongbuk. 712-714 South Korea

sjkang@daegu.ac.kr

² School of Computer Science & Engineering, Kyungshung University
Pusan. 608-736 South Korea

dbaisk@ks.ac.kr

³ Department of Computer Science, National University of Singapore
nash@comp.nus.edu.sg

Abstract. A hedge is a linguistic device used to avoid using a categorical sentence. Hedges can be used to determine whether a sentence is factual by merely regarding a sentence containing hedges as non-factual. In this paper, we perform a comparative experiment of various classification methods for hedge detection. Among four different classification methods, we observe that SVM shows the best performance and that the SVM-based method finally outperforms the best system in the CoNLL2010-ST task.

Keywords: Hedge Detection, Machine Learning, Natural Language Processing, Information Extraction.

1 Introduction

A hedge is a linguistic device used to avoid using a categorical sentence [1, 2]. Examples of hedges are *may*, *probably*, *it appears that*, etc. Hedges are used in a sentence when the writer is uncertain or has doubt about the contents of the sentence. Due to this uncertainty, sentences with hedges are considered to be speculative or non-factual.

Motivated by this characteristic of hedges, researchers have exploited hedges as useful features to extract factual sentences from documents. There are many applications which need to determine whether a sentence is factual or not. For example, in designing a specific type of question answering system which finds a factual answer, non-factual parts from documents need to be filtered out to effectively find an answer.

In this paper, we address the problem of hedge detection, i.e., to determine whether or not a sentence contains hedges. As hedge detection generally belongs to a classification problem, all supervised machine learning (ML) approaches can be applied to this problem. However, there has been no previous work on comparing existing methods, thus it is not clear what the best ML approach is for hedge detection. To draw a useful conclusion to this issue, we attempt to compare the state-of-the-art ML approaches for hedge detection. For the comparison, we selected CRF,

SVM, k -NN, and DT (decision tree learner), as they have been widely used for many classification problems. In our comparison results, SVM shows the best performance among four different classification methods and the SVM-based method finally outperforms the best system in CoNLL2010-ST task.

The paper is organized as follows. In Section 2, we summarize related work. In Section 3, we describe features selected to detect hedges. Experimental results are presented and discussed in Section 4. Finally, Section 5 puts forward some conclusions.

2 Relevant Work

Light et al. [3] used a simple technique, which determines hedge sentences relying on the presence of hedge cue words such as *suggest*, *potential*, *likely*, *may*, etc. There have been a few successful works on gathering a variety of hedge cue terms based on the weakly-supervised learning with SVM [4] or MaxEnt [5]. Such methods start from initial seeds such as *suggest* and *likely*, and extend the set of cue terms by exploiting the assumption that hedge cue words may co-occur with '*may suggest*'. Morante and Daelemans [6] reported 84.7% in F1 by using k -NN learning based on the features such as lemmas, words, POS, and IOB tags of its preceding and subsequent three tokens for each token in training sentences.

Recently, CoNLL 2010 Shared Task (CoNLL2010-ST)¹ evaluated the performance of detecting hedge cues and their linguistic scope in biological literatures and Wikipedia documents [7]. Its top-ranked systems adopted a sequence labeling (SL) approach [8-10] based on CRF or SVM_hmm.

Tang et al. proposed a cascaded system of CRF and SVM_hmm. However, the cascaded method was not better than that of their single CRF classifier, which achieved the best performance of 86.79% in the biological task of CoNLL2010-ST [8]. Tang et al. used words, lemmas, POS, chunk tags, and some composite features. Li et al. devised a greedy-forward feature selection scheme to obtain high precision, and showed F-measure of 85.89% using CRF [9]. Zhou et al. exploited the synonym features from WordNet, and showed F-measure of 86.32% based on CRF [10]. In summary, most top ranked systems employed CRF.

3 Features for Detecting Hedges

In order to fairly compare machine learning techniques, the same set of features should be used during the process of machine learning. The commonly-used features in the top ranked systems of CoNLL2010-ST [7] are given as follows.

- (1) Word Feature: $word(i)$ ($i = -2, -1, 0, +1, +2$)
- (2) Lemma Feature: $lemma(i)$ ($i = -2, -1, 0, +1, +2$)
- (3) Part-Of-Speech (POS) Feature: $POS(i)$ ($i = -2, -1, 0, +1, +2$)
- (4) Chunk Tag Feature: $chunk(i)$ ($i = -2, -1, 0, +1, +2$)

¹ <http://www.inf.u-szeged.hu/rgai/conll2010st/>

In the above, $word(0)$, $word(-n)$, and $word(+n)$ means respectively the current word, the n -th word to the left of the current word, and the n -th word to the right of the current word. $lemma(i)$ is the stemming result of $word(i)$. $POS(i)$ is the part-of-speech of $word(i)$. $chunk(i)$ is the chunk tag of $word(i)$, which are represented in the IOB2 format (B for BEGIN, I for INSIDE, and O for OUTSIDE) [11].

4 Experimental Results

As a test set to evaluate several ML approaches to hedge detection, this study uses the biological part of the CoNLL2010-ST dataset which contains biological abstracts and full articles from the BioScope (biomedical domain) corpus. Table 1 shows some statistics of the test set.

Similar to earlier works, this study adopted the following general steps for hedge sentence classification. First, for a given sentence, each token of the sentence is classified into hedge-class or non-hedge-class. The classes are represented in the

Table 1. CoNLL2010-ST biological dataset

	# of training sentences	# of testing sentences
Hedge	2,620	790
Non-hedge	11,921	4,213
Total	14,541	5,003

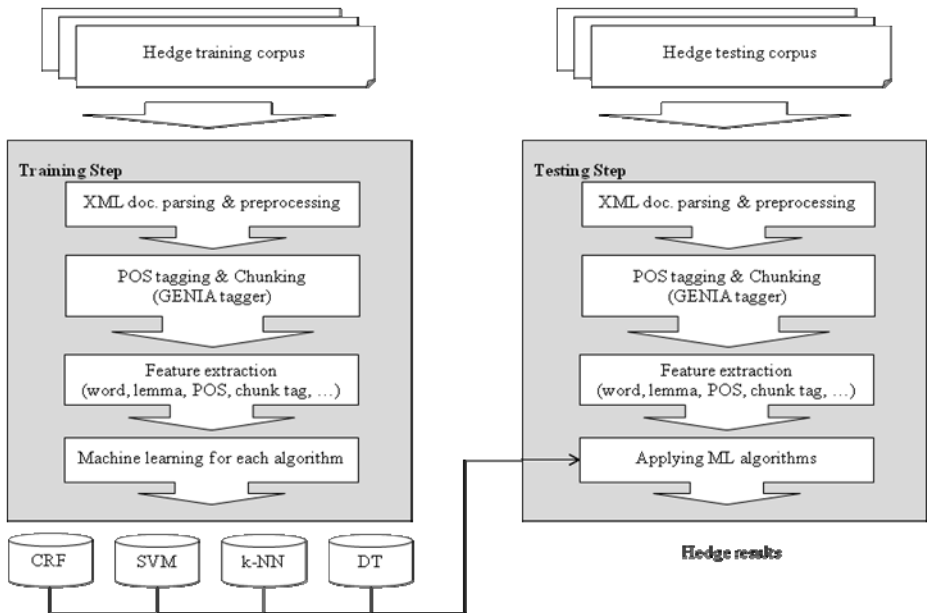


Fig. 1. The overall process for detecting hedges

IOB2 format (B for BEGIN, I for INSIDE, and O for OUTSIDE of hedge cue words) in order to annotate the cue phrases and the left and right boundaries of their scopes. Then, if there are one or more hedge-class tokens in a sentence, the sentence is classified into a hedge sentence.

The overall training and testing processes for detecting hedges are presented in Fig. 1. The training data has the information about locations of hedge cue words as well as whether each sentence is a hedge sentence or not. Basically, the training and testing step generate a list of raw sentences. POS tagging and chunking are performed against the raw sentences using the GENIA tagger which is known to work well on various types of biomedical documents [12]. From the output of GENIA tagger, the features defined in Section 3 are extracted, and then each classifier of CRF², SVM³, k -NN⁴, and DT⁵ is learned from the training data and evaluated using the test data.

This study evaluated two baseline methods. The first baseline system *Cue-dic* tries to detect hedge sentences solely relying on the cue-term list gathered from the train set. The second baseline system *Prob* finds hedge sentences based on the average hedge-probability of words in a sentence. Suppose that a sentence S consists of a sequence of words w_1, w_2, \dots, w_n . Then, the system *Prob* computes the hedge-probability $Score(S)$ of S which is defined using $Pr(w_i)$ the probability that word w_i appears in the hedge sentence. Then, the input sentence S is classified into the hedge class if $Score(S)$ is over a threshold.

$$Pr(w_i) = \frac{\text{frequency of } w_i \text{ in hedge sentences of training corpus}}{\text{frequency of } w_i \text{ in training corpus}} \quad (1)$$

$$Score(S) = \left(\sum_{i=1}^n Pr(w_i) \right) / n \quad (2)$$

Table 2 shows the evaluation results of six methods including four ML approaches which employ the same set of features defined in Section 3. *Cue-dic* used 1,344 hedge cue lists, which consist of 168 one-word cues, 420 two-word cues, and 756 more than three-word cues. In *Prob*, threshold 0.3 showed the best result among values between 0 and 1. For SVM, the 2nd degree of polynomial kernel and 1 slack variable were used. CRF used L2 for the regularization algorithm, 1.5 for hyper-parameter, and 1 for cut-off threshold for the features. In k -NN, when k is 2, k -NN showed the best result among k values between 1 and 4. DT used IB1 algorithm, weighted overlap for feature metrics, and GainRatio weight.

The result of *Cue-dic* indicates that a set of known hedge terms recalls most hedge sentences, but hedge cue terms appear nearly and equally in either hedge sentences or non-hedge ones. Interestingly, *Prob* method well discriminated hedge sentences, performing slightly better than or comparably to the two ML methods k -NN and DT.

² <http://crfpp.sourceforge.net/>

³ <http://chasen.org/~taku/software/yamcha/#rpm>

⁴ http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm

⁵ <http://www.rulequest.com/Personal/>

Table 2. The comparison results of hedge detection using the features defined in Section 3

Method		Precision	Recall	F1
Baseline	Cue-dic	48.04	94.56	63.71
	Prob	81.85	85.06	83.43
ML	CRF	87.14	84.05	85.57
	SVM	89.57	82.66	85.98
	<i>k</i> -NN	81.10	83.67	82.37
	DT	82.21	84.81	83.49

In later studies on detecting hedges, *Prob* could thus be regarded as a strong baseline approach.

Expectedly, SVM and CRF performed better than *k*-NN and DT, as reported in the previous studies. SVM and CRF showed greater precision than recall, while *k*-NN and DT made the opposite result. Unlike the results of the top-performing systems in CoNLL2010-ST, SVM achieved better effectiveness than CRF in this experiment, although the difference was marginal. Actually, two of the top-3 teams in CoNLL2010-ST attempted both HM-SVM and CRF to find the better. HM-SVM (Hidden Markov SVM) is a combination of HMM and SVM proposed for label sequence learning [8-9]. In our evaluation, however, SVM classifies each token in a sentence independently of other tokens. This is related to the difference between sequence labeling and token classification approaches to hedge detection. In other words, CRF and SVM in Table 2 can be viewed as respectively the representative methods of sequence labeling and token classification approaches. Sequence label learning like CRF may be more advantageous for hedge detection if the result of hedge detection is used by hedge scoping which should determine the boundaries of hedges in a sentence. Otherwise, however, token classification may be more appropriate in the sense that its learning process could focus on finding the most influential hedge token for improving sentence-level classification effectiveness instead of locating the entire hedge tokens of a single hedge as in sequence labeling.

We did additional experiments with SVM which performed best shown in Table 2. In addition to the common essential features defined in Section 3, we added *prev_res(j)* feature, which means the classification result of previous tokens within a window size. As shown in Table 3, SVM equipped with such additional feature led to improvement, yielding 86.82% in F1 as well as slightly outperforming the CoNLL2010-ST best systems. This indicates that sequential features like classification labels of preceding tokens may be helpful in detecting hedges using non-sequential labelers like SVM. As far as we know, previous tag feature like ours has not been successfully explored in hedge detection.

As to features for hedge sentence detection, the previous best systems employed not only well-known sentential contextual features (words, lemmas, POS, and chunks of neighboring words), but also additional complex features such as synonym

features, prefix/suffix, etc. For example, Tang [8] used several composite features such as a concatenation of the lemma and chunk of the same token. Unlike earlier state-of-the-art systems, the evaluation result of this study was obtained only from well-known local contextual features. This means that our result can be easily replicated by other researchers for further improvements.

Table 3. The results of hedge detection for different feature sets, using SVM for classifier

Feature set	Precision	Recall	F1
word(i), lemma(i), pos(i), chunk(i), prev_res(j) ($-1 \leq i \leq 1, -2 \leq j \leq -1$)	86.79	86.46	86.62
word(i), lemma(i), pos(i), chunk(i), prev_res(j) ($-2 \leq i \leq 2, -2 \leq j \leq -1$)	90.07	83.80	86.82
word(i), lemma(i), pos(i), chunk(i), prev_res(j) ($-3 \leq i \leq 3, -2 \leq j \leq -1$)	91.60	81.39	86.19
word(i), lemma(i), pos(i), chunk(i) ($-2 \leq i \leq 2$)	89.57	82.66	85.98
word(i), lemma(i), pos(i), chunk(i), prev_res(j) ($-2 \leq i \leq 2, j = -1$)	90.00	83.16	86.45
word(i), lemma(i), pos(i), chunk(i), prev_res(j) ($-2 \leq i \leq 2, -3 \leq j \leq -1$)	90.18	83.67	86.80

5 Conclusion

This study compared various machine learning techniques for detecting hedges and defined the feature set to achieve the best performance. Under the same condition, SVM and CRF showed better performances than k-NN and DT. In addition to common essential features such as words, lemmas, POS, and chunks, an additional sequential feature such as classification labels of previous words enabled a further improvement in SVM classifier. Fortunately, this study has found the best ML method and best set of features for hedge detection, which showed a slightly better effectiveness than the previous state-of-the-art hedge detectors.

In the future, we will apply our SVM-based hedge detector to the query answering task and examine the effect of using a hedge detector on the performance of answer extraction.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0007025).

References

1. Lakoff, G.: Hedges: a study in meaning criteria and the logic of fuzzy concepts. *Chicago Linguistics Society Papers* 8, 183–228 (1972)
2. Hyland, K.: Persuasion, interaction and the construction of knowledge: representing self and others in research writing. *International Journal of English Studies* 8(2), 8–18 (2008)
3. Light, M., Qiu, X.Y., Srinivasan, P.: The language of bioscience: facts, speculations, and statements in between. In: *Proceedings of BioLINK 2004: Linking Biological Literature, Ontologies and Databases*, pp. 17–24 (2004)
4. Medlock, B., Briscoe, T.: Weakly supervised learning for hedge classification in scientific literature. In: *Proceedings of 45th Meeting of the Association for Computational Linguistics*, pp. 992–999 (2007)
5. Szarvas, G.: Hedge classification in biomedical texts with a weakly supervised selection of keywords. In: *Proceedings of 46th Meeting of the Association for Computational Linguistics*, pp. 281–289 (2008)
6. Morante, R., Daelemans, W.: Learning the scope of hedge cues in biomedical texts. In: *Proceedings of the BioNLP 2009 Workshop*, pp. 28–36 (2009)
7. Farkas, R., Vincze, V., Mora, G., Csirik, J., Szarvas, G.: The CoNLL 2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text. In: *Proceedings of the Shared Task, 14th Conference on Computational Natural Language Learning, Sweden*, pp. 1–12 (2010)
8. Tang, B., Wang, X., Yuan, B., Fan, S.: A Cascade Method for Detecting Hedges and their Scope in Natural Language Text. In: *Proceedings of the Shared Task, 14th Conference on Computational Natural Language Learning, Sweden*, pp. 13–17 (2010)
9. Li, X., Shen, J., Gao, X., Wang, X.: Exploiting Rich Features for Detecting Hedges and Their Scope. In: *Proceedings of the Shared Task, 14th Conference on Computational Natural Language Learning, Sweden*, pp. 78–83 (2010)
10. Zhou, H., Li, X., Huang, D., Li, Z., Yang, Y.: Exploiting Multi-Features to Detect Hedges and Their Scope in Biomedical Texts. In: *Proceedings of the Shared Task, 14th Conference on Computational Natural Language Learning, Sweden*, pp. 106–113 (2010)
11. Sang, T.K., Veenstra, J.: Representing Text Chunks. In: *Proc. of EACL 1999*, pp. 173–179 (1999)
12. Tsuruoka, Y., Tateishi, Y., Kim, J.-D., Ohta, T., McNaught, J., Ananiadou, S., Tsujii, J.: Developing a Robust Part-of-Speech Tagger for Biomedical Text. In: Bozanis, P., Houstis, E.N. (eds.) *PCI 2005. LNCS*, vol. 3746, pp. 382–392. Springer, Heidelberg (2005)

Application of Trend Detection of Technical Terms to Technology Opportunity Discovery

Mi-Nyeong Hwang, Min-Hee Cho, Myung-Gwon Hwang,
Mikyong Lee, and Do-Heon Jeong

Dept. of Information Technology Research, KISTI, 335 Gwahangno,
Yuseong, Daejeon, 305-806, Korea
{mnhwang, mini, mgh, jerryis, heon}@kisti.re.kr

Abstract. This paper proposes a new approach to trend detections for technical terms extracted from academic literature and also describes its application. Because the trends of technical terms can express the changes of particular subjects in a specific research field over time, it can be successfully utilized for researchers to establish future research plan. The experiments shown in this paper is performed by exploiting the NDSL (National Discovery for Science Leaders) academic literature data maintained by KISTI. This paper also introduces how the proposed method can be applied to InSciTe, a service that supports technology opportunity discovery (TOD).

Keywords: Technical Term, Term Usage Cycle, Trend Detection, Trend Analysis.

1 Introduction

Since defined by Tim Berners-Lee in the late 1990's, the semantic web has made continuous development. The current semantic made it possible for agents to detect trends¹ and are providing users with the meanings extracted and inferred from the data on heterogeneous platforms².

As the semantic web is growing and expanding, the techniques for identifying, extracting, and analyzing the technical entities has become the linchpins of the technical documents analysis. Among those techniques, trends analysis and topic detection and tracking (TDT) are currently attracting academic and industrial attention. The trend, or the tendency or inclination along the time sequence, can be detected through the appearance patterns of entities or concepts expressed as noun phrases in a certain document corpus. In other words, we can detect and predict the technology trends by extracting and analyzing the technical entities from scientific and technical documents.

A new analysis method for detecting the research trends in academic papers is proposed in this paper. In order to implement the proposed method, we extracted

¹ FUSE(Foresight and Understanding from Scientific Exposition).

² CUBIST(Combining and Uniting Business Intelligence with Semantic Technologies).

technical terms from research paper database, and periodically evaluated the term dominance value (TDV) of each term within specific time duration. Then the each term has its own time series TDV data, and we classified the terms with the same patterns into a single trend group.

2 Related Work

Emerging trend detection (ETD) is a challenging problem on which research attention is intensively focused, and its goal is to detect newly appearing concepts or entities based on the frequency of words [1].

There are a few commercial services. ‘Google trends’³ visualizes the frequency change of the query keywords. ‘Google topic’⁴ service does not only inform users of the frequent keywords, but also detects the hot topics dealt in news, blog, and other services. ‘BlogPulse’ detects the trends based on the ‘burstiness’ defined with the 2-week average frequency and daily frequency of each noun phrase extracted from web blogs [2]. ‘Trendsmap’⁵ provides real-time information on the world-wide trends of twitter messages.

There have been also a few approaches to the detection of researchers’ interest in a certain academic area. Abe and Tsumoto tried to detect the trends by periodically computing the ‘*term frequency-inverse document frequency* (TF-IDF)’ from the corpus with time-series information [3].

In order to analyze the usage cycle of each term, Jung *et al.* proposed ‘*term dominance value* (TDV)’ computed with the appearance frequency and different weights on each observation period [4]. Taking TDV into account, they classified the current state of each term into different groups, ‘*creation*’, ‘*growth*’, ‘*declination*’, and ‘*extinction*’. However, the final trend analysis of the method was still limited to distinguish ‘*emergent*’ terms and ‘*subsiding*’ ones.

In this paper, we employed the trend analysis method that models the usage cycle of technical terms by classifying the terms based on the patterns of their time series TDV data [7]. This paper also introduces the application of the method to InSciTe, an actual service for discovering the technology opportunity based on trend analysis.

3 Technical Term Usage Cycle Modeling

In this section, the process of term usage cycle modeling is described in briefly. The technical details have been already elaborated in the previous work [7]. The framework of the process is shown in Figure 1. In the first step, technical terms are extracted from the large document corpus with temporal information attached, and TDV of each term is computed. Then, the time series patterns of TDV are computed, and the terms are divided into several groups based on the patterns. The final trend analysis is performed in each group. The steps are listed as follows:

³ Google Trends, <http://www.google.com/trends>

⁴ Google Topic, <http://www.google.co.kr/topicsearch>

⁵ Trendsmap, <http://trendsmap.com>

1. Technical term extraction - identification and extraction of terms from corpus.
2. TDV calculation - computation of TDV of each term for each time period.
3. Pattern clustering-classification of terms based on the temporal patterns of TDV
4. Trends detection - assignment of meanings to trend groups.

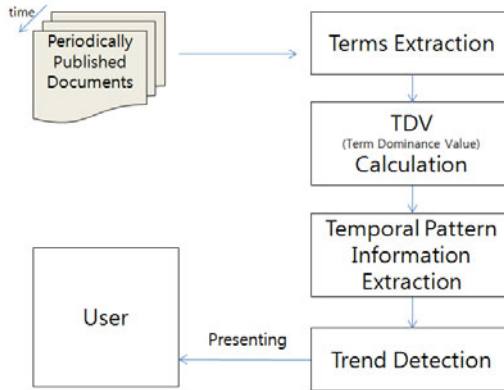


Fig. 1. Overview of the term usage cycle modeling

In the first step of the term usage cycle modeling, technical terms are extracted from a large document corpus. In order to extract the terms, SINDI-CORE (Scientific Intelligent Discovery), a text mining system introduced in [5] was exploited.

In order to detect the trends of technical terms appearing in scientific and technical literature, we applied the propose method to academic papers stored in NDSL (National Discovery for Science Leaders, <http://www.ndsl.kr>) database⁶ from 2006 to 2011. The NDSL provides information search services with the easy and versatile queries on various academic documents such as papers, patents, and trends reports. However, the service is not well-suited for text mining or research trends analysis.

The experimental data is composed of technical terms extracted from the academic papers in the NDSL database. The subject category of each paper was also taken into account according to ‘National Standard Science and Technology Classification’, and the same terms in different categories were considered different entities. The term extraction process identified and normalized technical terms from the abstracts of the documents by employing SINDI, a text mining system. After the extraction, only the terms that appeared more than 10 times and also appear in Wikipedia service were selected as technical terms. The number of total technical terms selected was 48,775. The usage cycle modeling was separately applied to each subject category, and terms were classified into several trend groups such as ‘growing’, ‘growth-continuing’, ‘slowing-down’, ‘declining’, ‘extinctive’, ‘maintaining’, and ‘reviving’ groups within each category.

⁶ NDSL paper data is as of March 2011.

4 Application to Technology Opportunity Discovery

Clustering TDV-based time series patterns enables researchers to detect the trends of technical terms from a specific time to another specific time. By means of the trends detection, it is possible for the researchers to have the insight into the scientific activities in the field. It is also possible for the researchers to understand the newly emerging issues, enduring topics, and more important research items, and to select their future topics for the successful research. By considering, for instance, the terms in the ‘reviving’ group such as ‘distributed computing environment’, ‘distributed generation’, and ‘mobile device’ in the IT research field, we can notice that the concept of ‘cloud computing’ related with those terms is gaining more attention in academic and industrial areas. Such knowledge is crucial for the researcher and entrepreneurs to make a successful research and development of future technology.

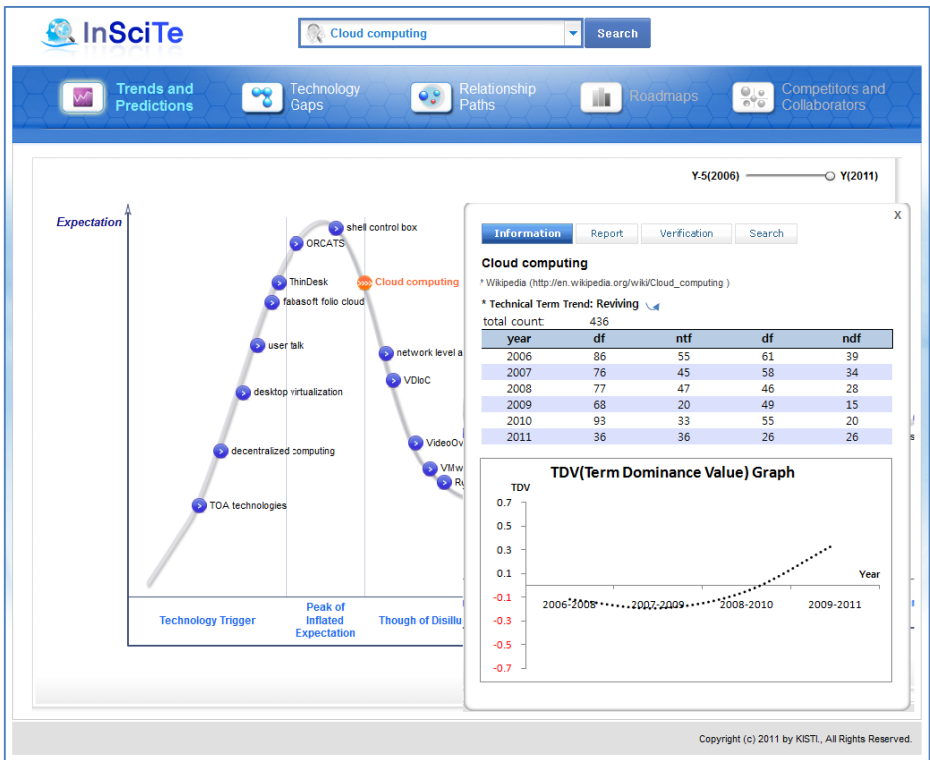


Fig. 2. InSciTe Service Screenshot: tooltip on ‘Cloud computing’ visualizes the technical term trend based on the usage cycle model

The method introduced in this paper was applied to InSciTe, a technology opportunity discovery (TOD) service that supports technical decision-making based on the technology trends. This service visualizes the current life cycles of emerging

and related technologies by mapping them on Hype Cycle graph⁷. If a user searches a technology of interest, the growth or maturity of the technology is visualized in InSciTe. Moreover, information on the similar or related technologies can be also obtained, and users can find which research institutes or countries are intensively studying and developing such technologies. If detailed information on those technologies is needed, their trends within the academic literature are detected by the method described in this paper. As shown in figure 2, the quantitative data such as annual term-frequency, document-frequency, and TDV graph are visually provided.

5 Conclusion

In this paper, we described a method that detects the trends of technical terms from academic literature, and also introduced the application of the method to an actual TOD system. Although the previous approach to trend detection generally divided the trends into two major groups, ‘*emerging*’ and ‘*declining*’ groups, the proposed method subdivides the trends into 7 groups (i.e., ‘*growing*’, ‘*growth-continuing*’, ‘*slowing down*’, ‘*declining*’, ‘*extinctive*’, ‘*maintaining*’, and ‘*reviving*’). The method can effectively help researcher to make a proper decision on future R&D strategies, and will be fully integrated into a new technology intelligence service such as InSciTe.

In the further research, authors will investigate whether different aspects should be considered when the trend detection method is applied to different research fields.

References

1. Kontostathis, A., Galitsky, L.M., Pottenger, W.M., Roy, S., Phelps, D.J.: A Survey of emerging trend detection in textual data mining. In: Survey of Text Mining: Clustering, Classification, and Retrieval, pp. 185–224 (2003)
2. Gance, N.S., Hurst, M., Tomokiyo, T.: BlogPulse: Automated Trend Discovery for Weblogs. In: WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics (2004)
3. Abe, H., Tsumoto, S.: Analysis of Research Keys as Temporal Patterns of Technical Term Usages in Bibliographical Data. In: An, A., Lingras, P., Petty, S., Huang, R. (eds.) AMT 2010. LNCS, vol. 6335, pp. 150–157. Springer, Heidelberg (2010)
4. Jung, H., Koo, H., Lee, B., Sung, W.: Study of Term Life Cycle Management for Efficient Resource Use. In: KCC, vol. 32(1), pp. 45–459 (2005) (in Korean)
5. Choi, Y., Jeong, C., Cho, H.: A Study on the Integration of Recognition Technology for Scientific Core Entities. *Journal of Information Management* 28(1), 89–104 (2011)
6. Lee, M., Jung, H., Kim, P., Sung, W.: Technology Intelligence Service for Establishing R&D Strategy. *Journal of KIISE* 17(5), 337–341 (2011) (in Korean)
7. Hwang, M., Cho, M., Hwang, M., Jeong, D.: Trend Analysis of Technical Terms Using Term Life Cycle Modeling. *Journal of KIPS* (submitted, 2011) (in Korean)

⁷ Gartner’s Hype Cycle: <http://www.gartner.com/technology/research/hype-cycles/>

Decision-Making Support Service Based on Technology Opportunity Discovery Model

Mikyoung Lee¹, Seungwoo Lee¹, Jinhyung Kim¹, Dongmin Seo¹,
Pyung Kim¹, Hanmin Jung¹, Jinhee Lee^{1,2}, Taehong Kim^{1,2},
Hee Kwan Koo^{1,2}, and Won-Kyung Sung¹

¹ Korea Institute of Science and Technology Information (KISTI),
335 Gwahngno, Yuseong-gu, Daejeon, Korea 305-806

² University of Science and Technology (UST),
113 Gwahangno, Yuseong-gu, Daejeon, Korea 305-333
{jerryis, swlee, jinhyung, dmseo, pyung, jhm, jhlee29,
kimtaehong, hkkoo, wksung}@kisti.re.kr

Abstract. This paper describes InSciTe Advanced, a decision-making support service, based on TOD(technology opportunity discovery) model. TOD model is a logical model for discovery of emerging technologies and prediction of phase and speed on a technology life cycle. InSciTe Advanced is based on semantic technologies such as ontology, semantic repository and inference as well as text mining. It aims to provide multi-facet services on emerging technologies, their elements and alternations in all domain. InSciTe Advanced has major services such as trends and predictions, technology levels, relationship paths, roadmaps and competitors and collaborators.

Keywords: Technology Opportunity Discovery, Semantic Technologies, Decision-Making Service.

1 Introduction

Today, global technology competitiveness comes from discovering technology opportunity and dominating the technology in advance. Analyzing technology opportunity and reflecting it in R&D activities have been emphasized on securing R&D competitiveness in the nations and companies.

Technology Opportunity Discovery (TOD) is activities that enable companies to identify the technological opportunities and threats that could affect the future growth and survival of their business. It aims to capture and disseminate information about emerging technologies and trends needed for strategic planning and decision making.

Researchers can draw future technologies and business items in advance from TOD. TOD also helps save time and efforts for finding reasonable information and contributes to improve R&D efficiency. In addition to, it leads to strengthen the technology competitiveness of researcher and small and medium-sized business and also induces the high technology commercialization.

In this paper, we explain a decision-making support service named InSciTe Advanced which is able to improve the global technology competitiveness of researchers and companies by providing emerging technologies under the analysis of mass academic literature.

2 Related Work

In the early of 1990s, Georgia Tech's Alan Porter started to do research on framework about Technology Opportunity Analysis(TOA) and aimed at systematizing the process that prioritizes R&D investment in the emerging technology areas[1]. In 2006, VantagePoint provided a study on discovering R&D direction, key technology areas and main researchers using text-mining technologies[2].

Since 2010, IARPA(Intelligence Advanced Research Projects Activity) in U.S. has planed FUSE (Foresight and Understanding from Scientific Exposition) project. FUSE approach is to combine explicit metadata used by most services with implicit metadata hidden within text documents, in order to effectively find technical emergence. The FUSE Program seeks to develop automated methods that aid in the systematic, continuous, and comprehensive assessment of technical emergence using information extracted from the published scientific, technical, and patent literature[3].

These studies commonly aim to make automated system that supports for researchers and firms to find emerging technologies.

So we developed a decision-making support system that discovers emerging technologies and new business items by analyzing mass data automatically without expert's assistance.

3 Service

InSciTe Advanced is a decision-making support service based on technology opportunity discovery model. InSciTe Advanced extracts meaningful technologies and their relations from large-scale academic literature and combines the results with meta-data on a semantic service platform to enhance their analytical values. InSciTe Advanced is based on semantic technologies such as ontology, semantic repository and inference as well as text mining.

InSciTe Advanced's goal is to support researchers of institutes and companies who want to find emerging technologies and technology trends using in-depth technology analysis. To cope with the goal, we designed the TOD model which is able to discover future emerging technologies automatically and developed visualization service. TOD model is a logical model for discovery of emerging technologies and prediction of phase and speed on a technology life cycle. The former is done at emerging technology discovery model and the latter is done at technology life cycle discovery model[4].

InSciTe Advanced also provides composite services developed by several different viewpoints on the relation such as technology-technology, technology-agent, agent-technology and agent-agent.

Specific services developed by using various viewpoints include the following:

Trends and predictions that shows the relations among technologies using technology life cycle(TLC); *Technology levels* that represents technology level and technology penetration time of research agents; *Relationship paths* that displays the relations between technology and agents; *Roadmaps* that expresses technology trends and candidate technologies for research forecasting; *Competitors and collaborators* that shows information about competition or cooperation among agents

● **Trends and predictions**

Trends and Predictions service shows development stages of technology on the Hype cycle under TOD model base. It helps researchers to predict each technology development stage by year and forecast technology development speed.

This service, which analyzes the relations among technologies such as similar technology, element technology and concurrent technology, represents technology life cycle defined under the TOD model in Hype cycle graph. It provides element technologies, and similar technologies using text mining technology and related technologies according to the Wikipedia’s category.

As shown in Fig. 1, users can identify technology life cycle of the searched and related technologies.

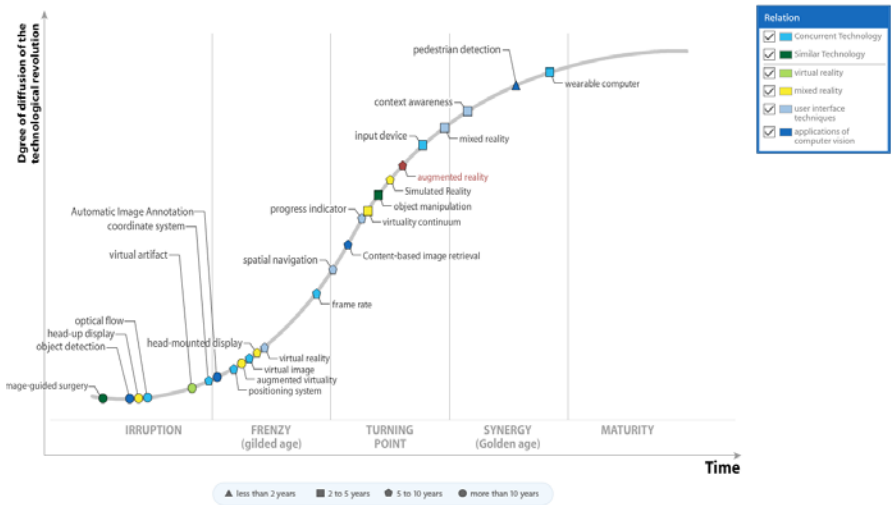


Fig. 1. Trends and predictions

● **Technology levels**

This service shows current technology level and entry point of country or institution about searched technology. Color of the node means cooperative research groups of agents, and size of the node displays R&D results of agents. Technology level indicates the gap of technology level between the leading agent and other agents, and new entry displays the point of R&D results from the first time.

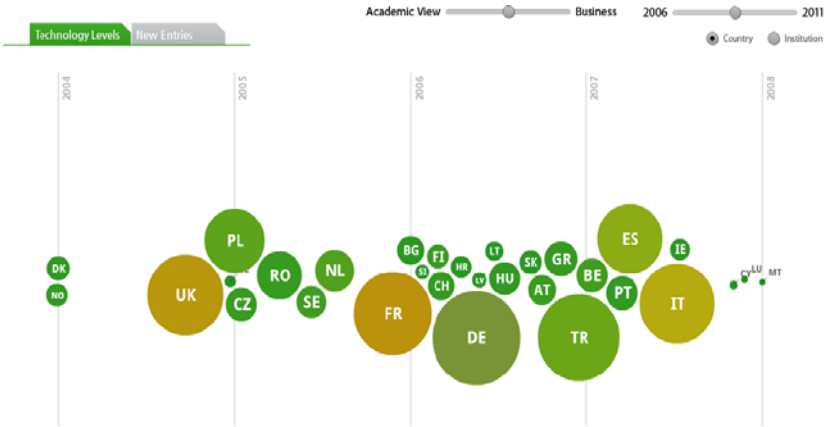


Fig. 2. Technology Levels

The main feature of this service is able to display a variety of information about R&D cooperation information of the agents, leading research agents, technology gap among agents and new entry of technology.

● **Relationship Paths**

Relationship Paths is a service that traces the correlation between technology and agent by using OntoRelFinder. OntoRelFinder is an inter-instance path-finding engine that fast traverses and visualizes all possible paths among more than two instances from large-scale ontology database[5].

It is possible to find the relationship of the technologies and agents, because technology and agent is class and their relation is property on the Ontology Model. Users know the direct and indirect relationship of the technologies and agents from this service. Fig. 3 shows the relationship paths between ‘cloud computing’ and ‘virtual reality’.

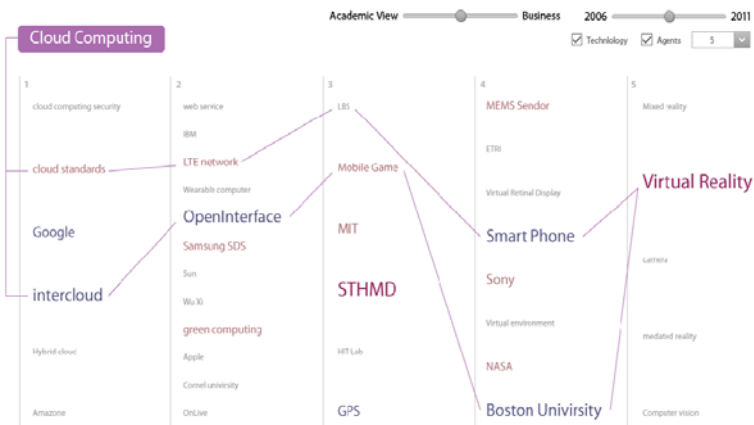


Fig. 3. Relationship Paths

● **Roadmaps**

Technology roadmaps service is to propose appropriate technologies related with the technologies studied by an agent in the near future.

We designed an algorithm of candidate technologies discovery in order to propose the emerging technology while considering technology trend analysis of leading agents and technology development process analysis.

From customer’s perspective, this roadmap service can be used on R&D planning and management. Fig. 4 shows major technologies trends and candidate technologies for research forecasting.

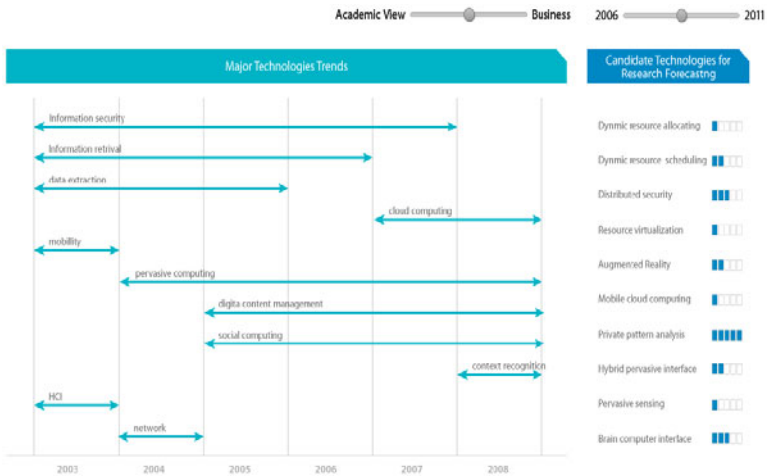


Fig. 4. Roadmaps



Fig. 5. Competitors and Collaborators

● Competitors and collaborators

Competitors and collaborators service shows competition and collaboration relations among agents that conduct similar research. Researchers know that a searched agent cooperates with which agents on what technology. Firms and researchers can set R&D strategy directions through comparing the competition and collaboration level among agents conducting same research topic.

As shown Fig. 5, this service displays same research technology lists of the agent and represents cooperative, competitive and similar agent lists.

4 Conclusion

In this paper, we explain the InSciTe Advanced service that supports decision making for strategy establishment by applying semantic technologies to mass literatures in science and technology. InSciTe Advanced is focusing on discovering emerging technologies by using the TOD model. The advantage of InSciTe Advanced is able to provide TOD information about the relations of technology and agent to researchers. InSciTe Advanced provides 5 main services; Trends and Predications, Technology Levels, Relationship Paths, Roadmaps, Competitors and Collaborators.

In the future, we will add automatic summary reporting service and infometrics service.

References

1. Porter, A.L., Detampel, M.J.: Technology opportunities analysis. *Technological Forecasting and Social Change* 49(3), 237–255 (1995)
2. Richard, G.: Scaling the technology opportunity analysis text data mining methodology. Data extraction, cleaning, online analytical processing analysis and reporting of large multi-source datasets. Capella University. Doctoral Dissertation (2006) ISBN: 978-0-542-81605-5
3. http://www.iarpa.gov/solicitations_fuse.html
4. Lee, J., Kim, J., Jung, H., Sung, W.-K.: Toward Discovering Emerging Technologies Based on Decision Tree. In: *IEEE CPSCOM 2011* (accepted, 2011)
5. Seo, D., Jung, H., Kim, P., Lee, S., Lee, M., Sung, W.-K.: Finding Relations among Entities using Ontology Schema Paths. *Proc. Korean Society for Internet Information* 12(1), 331–332 (2011) (in Korean)

Relation Extraction Based on Composite Kernel Combining Pattern Similarity of Predicate-Argument Structure

Hong-Woo Chun, Chang-Hoo Jeong, Sa-Kwang Song,
Yun-Soo Choi, Sung-Pil Choi, and Won-Kyung Sung

Korea Institute of Science and Technology Information, 245 Daehangno, Yuseong-gu,
Daejeon, 305-806, South Korea

{hw.chun, chjeong, esmallj, armian, spchoi, wksung}@kisti.re.kr

Abstract. Lots of valuable textual information is used to extract relations between named entities from literature. Composite kernel approach is proposed in this paper. The composite kernel approach calculates similarities based on the following information: (1) Phrase structure in convolution parse tree kernel that has shown encouraging results. (2) Predicate-argument structure patterns. In other words, the approach deals with syntactic structure as well as semantic structure using a reciprocal method. The proposed approach was evaluated using various types of test collections and it showed the better performance compared with those of previous approach using only information from syntactic structures. In addition, it showed the better performance than those of the state of the art approach.

1 Introduction

Extraction of relations between meaningful entities from literature is an encouraging research topic in the field of information extraction[1, 2, 3]. Various contextual information is used as features to extract relations. This paper presents a method using the following two types of features: (1) Phrase structure in convolution parse tree kernel. (2) Predicate-argument structure (PAS) patterns.

2 Previous Research

In the fields of relation extraction, kernel-based methods have shown encouraging results. Zhang et al. (2006) proposed a composite kernel approach that uses various syntactic structural information and entity features provided by ACE corpus[4]. In this research, they obtained various syntactic structural information using the convolution parse tree kernels proposed by Collins and Duffy[5].

Guo-Dong et al. (2007) showed a parse tree kernel-based approach that applies a tree pruning technique and context sensitive structured parse tree information, and this approach can deal with contextual information with respect to entities[6].

We propose a composite kernel approach that combines the parse tree kernel and the PAS pattern-based similarity kernel. The proposed approach showed promising evaluation results.

3 Composite Kernel-Based Relation Extraction

3.1 Parse Tree Kernel-Based Relation Extraction

The main idea of a parse tree kernel is to sever a parse tree into its sub-trees and transfer it as a point in a vector space in which each axis denotes a particular sub-tree in the entire set of parse trees. If this set contains M unique sub-trees, the vector space becomes M dimensional. The similarity between two parse trees can be obtained by computing the inner product of the two corresponding vectors, which is the output of the parse tree kernel. To optimize the parse tree kernel, we applied a tree pruning technique. Path-enclosed tree (PT) pruning method is the current state of the art method[7], so we used the PT method and dealt with sub-trees constructed from the shortest syntactic path between two entities. Figure 1 describes an example of a sub-tree constructed by the pruning method.

3.2 Predicate-Argument Structure Pattern-Baed Relation Extraction

Predicate-argument structure (PAS) can analyze predicate-argument relations among all words in a sentence, and a PAS pattern is the shortest PAS path between candidate entities. Figure 2 describes PAS of a sentence and a PAS pattern for two entities: *Radon exposure* and *lung cancer*.

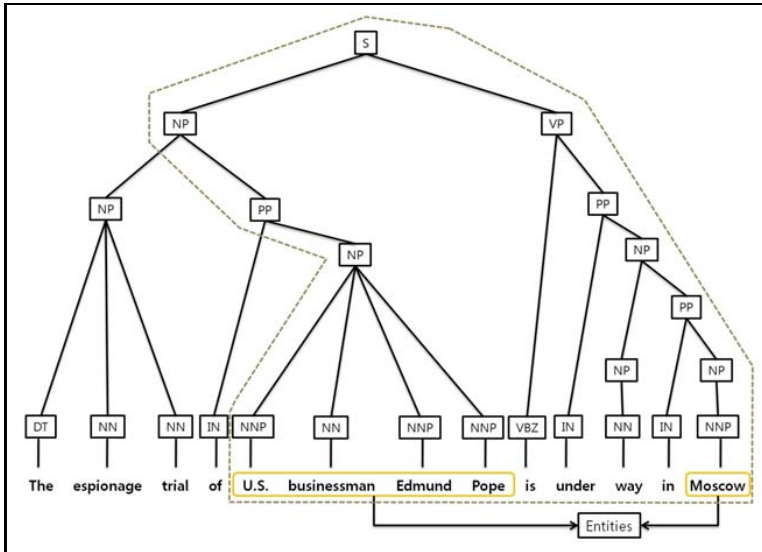


Fig. 1. An example of sub-tree constructed by the PT pruning

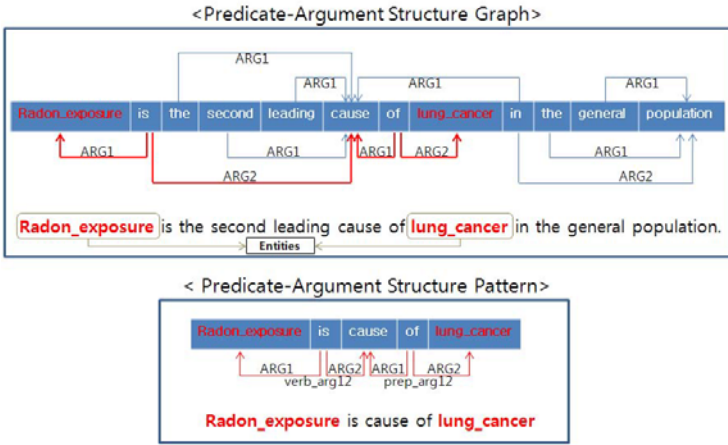


Fig. 2. Predicate-argument structure-based pattern generation

3.3 Construction of Composite Kernel

Composite Kernel approach combines the parse tree kernel and PAS pattern-based kernel linearly (Equation 1).

$$K_{composite} = (1 - \tau) \times K_{tree}(I_1, I_2) + \tau \times K_{pas}(I_1, I_2) \quad (1)$$

In Equation 1, $K_{tree}(I_1, I_2)$ is a parse tree kernel that calculates similarity of phrase structures of I_1 and I_2 , and $K_{pas}(I_1, I_2)$ is a kernel that calculates similarity of PAS patterns of I_1 and I_2 . τ plays a role in optimizing weight.

In the proposed approach we used several open techniques: (1) A parse tree kernel method developed by Moschitti et al. [8] is applied in the proposed method to calculate the parse tree kernel-based similarity quickly. (2) LIBSVM [1] is used for training Support Vector Machines. (3) Charniak parser [2] is used to construct phrase structured trees. (4) ENJU parser [3] is used to analyze PAS patterns.

4 Experiments

The candidate entities in the proposed approach are *Person*, *Location*, *Organization*, and *Terminology (PLOT)*. For evaluating and training the relation extraction system, KREC 2010 corpus is constructed. KREC 2010 corpus was collected with respect to Green Technology. There are following three criteria in collection policy: (1) As for the journal articles, the impact factor is a criteria. We selected highly ranked 50 journals for the collection. (2) The average length of articles was calculated, and articles were selected if their length were longer

¹ <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

² <http://www.cs.brown.edu/~ec/#software/>

³ <http://www-tsujii.is.s.u-tokyo.ac.jp/enju/>

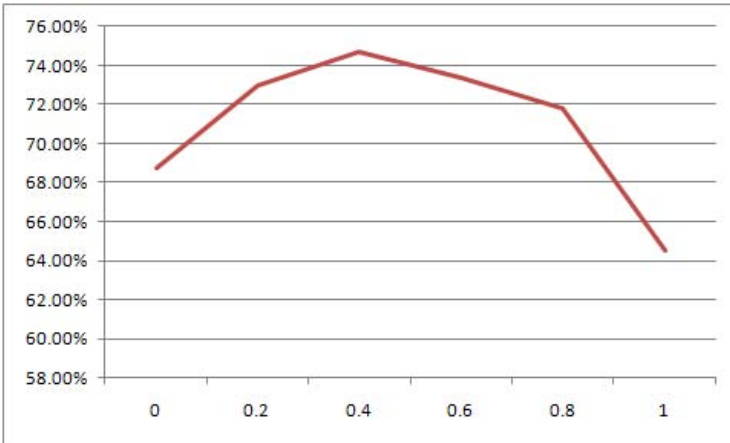


Fig. 3. Evaluation results based on τ

than 90% of the average length. (3) Publication year was a criteria. Articles published from 2000 were selected.

Table 1 describes statistics of KREC 2010. The number of relation classes in KREC 2010 corpus is 39, but 7 of them are used in the evaluation: *relate*, *change*, *produce*, *own*, *connect*, *analyze*, and *cause*.

Table 1. Statistics of KREC 2010

# Documents	# Sentences	# Entities	# Relations
1,090	14,341	22,125	2,441

Optimized value τ is decided as 0.4 through the experiments (Figure 3).

Table 2. Performances

Types	mi-F1(%)	ma-F1(%)
PAS pattern-based similarity kernel	64.60	33.69
Parse tree kernel	68.78	38.09
Composite kernel	74.72	42.33

Notes: mi-F1: micro-avaeraged F-score, ma-F1: macro-avaeraged F-score.

Table 2 presents the experimental results. As Table 2 describes, the composite kernel-based method shows the best performance, and the parse tree kernel-based method shows the better performance than those of PAS pattern similarity kernel-based method. We found that the PAS pattern-based similarity information plays an import role in improving performance of the parse tree kernel-based approach.

5 Conclusion

The composite kernel is proposed to extract relations among Persons, Locations, Organizations, and Terminologies (PLOT). The composite kernel-based model is constructed by combining a parse tree kernel-based model with a predicate-argument structure (PAS) pattern-based similarity. Through several experiments, we found that PAS pattern-based similarity information plays an important role in improving the performance.

Because the proposed approach has been developed independently to domains, it might be applied to various types of corpora.

References

1. Bunescu, R.C., Mooney, R.J.: A Shortest Path Dependency Kernel for Relation Extraction. In: Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Vancouver, B.C, pp. 724–731 (2005)
2. Culotta, A., Sorensen, J.: Dependency Tree Kernels for Relation Extraction. In: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (2004)
3. Bunescu, R.C., Mooney, R.J.: Subsequence Kernels for Relation Extraction. In: Advances in Neural Information Processing Systems (2006)
4. Zhang, M., Zhang, J., Su, J., Zhou, G.: A Composite Kernel to Extract Relations between Entities with both Flat and Structured Features. In: 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pp. 825–832 (2006)
5. Collins, M., Duffy, N.: Convolution Kernels for Natural Language. In: NIPS 2001 (2001)
6. GuoDong, Z., Min, Z., Dong, H.J., QiaoMing, Z.: Tree Kernel-based Relation Extraction with Context-Sensitive Structured Parse Tree Information. In: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Prague, pp. 728–736 (June 2007)
7. Zhang, M., GuoDong, Z., Aiti, A.: Exploring syntactic structured features over parse trees for relation extraction using kernel methods. *Information Processing and Management* 44, 687–701 (2008)
8. Moschitti, A.: Efficient Convolution Kernels for Dependency and Constituent Syntactic Trees. In: Fürnkranz, J., Scheffer, T., Spiliopoulou, M. (eds.) ECML 2006. LNCS (LNAI), vol. 4212, pp. 318–329. Springer, Heidelberg (2006)

English-to-Korean Cross-Lingual Link Detection for Wikipedia

Ralph Marigomen and In-Su Kang

School of Computer Information, Kyungsook University, 314-79, Daeyon-dong, Nam-gu,
Busan 608-736, South Korea
rmarigomen@gmail.com, dbaisk@ks.ac.kr

Abstract. In this paper, we introduce a method for automatically discovering possible links between documents in different languages. We utilized the large collection of articles in Wikipedia as our resource for keyword extraction, word sense disambiguation and in creating a bilingual dictionary. Our system runs using these set of methods for which given an English text or input document, it automatically determines important words or phrases within the context and links it to a corresponding Wikipedia article in other languages. In this system we use the Korean Wikipedia corpus as the linking document.

Keywords: Wikipedia, keyword extraction, word sense disambiguation.

1 Introduction

Annotating a document tends to provide users an easy way of accessing additional information regarding a particular topic. Content providers used to manually annotate important keywords or key phrases which gives higher level of explanation for their content. These keywords are linked to a certain or related articles providing a more elaborative description.

The Cross-lingual Link Detection (CLLD) is a method for automatically determining these keywords or key phrases from a given context or document and links it to a corresponding document with different languages [1]. It allows readers to have access to documents in different languages which could either have a richer context or something they are more familiar with thus eliminating the language barrier in information sharing. Fig. 1 shows an illustration of CLLD.

Wikipedia is a free online multilingual encyclopedia consisting of a vast number of articles resulting from the collaborative efforts of many volunteer contributors. It has been extensively used as a resource in many natural language processing tasks such as named entity disambiguation, keyword extraction, text summarization, etc. Wikipedia editions are available in more than 200 languages with English having more than 3 million articles.

The CLLD system is composed of three core functions, namely, keyword extraction, translation, and word sense disambiguation. Discussing briefly, it first retrieves candidate keywords from any given input document. It then translates each of these candidate terms into Korean and linked them to its corresponding Korean

Wikipedia article. In this system, we utilized the large collection of documents from Wikipedia as our resources for performing these three core functions of the system. The extensive hypertext links or manually annotated keywords within a Wikipedia document can be used as a tool in keyword extraction and can also be used for evaluating such methods. Equivalent documents of some topics written in different language are useful in providing a list of bilingual pairs such as English to Korean pairs which are helpful for translation processes. Our system is not directly related to any language translation approach. An approach, such as machine translation, performs by directly converting or substituting sentences into its equivalent meaning for a target language. CLLD system on the other hand actively identifies meaningful words or phrases in a given context and links them to a corresponding article, thus giving such terms more definitive information in another language.

The CLLD can be beneficial for content providers who want to provide their readers contents enriched with Korean encyclopedic information.

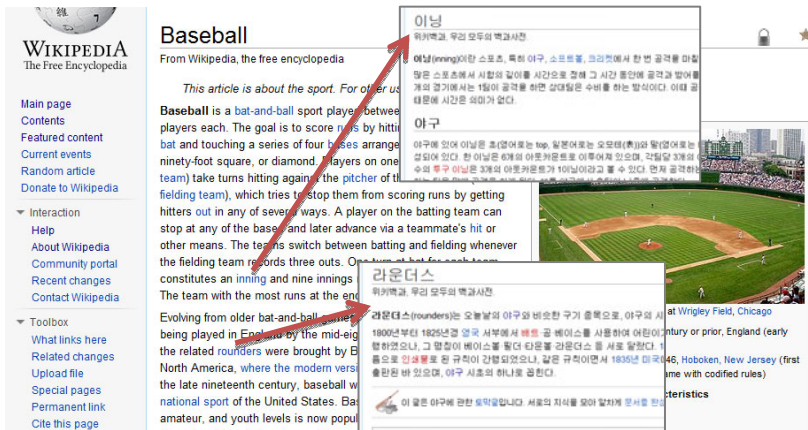


Fig. 1. An illustration of the CLLD system which links important keywords in an English document to its corresponding Korean Wikipedia article

2 Related Work

There is a lot of existing link discovery tools that can assist content providers in determining potential anchors and targets from a given content. But there are no existing tools that can support the linking of documents from different languages. One approach for such key term extraction is a so called Communities-based [2] approach which models a document as a graph of interconnected term based on their semantic relationship. A work closely related to our CLLD system is the “Learning to Link with Wikipedia” by Milne and Witten [3] and the “Wikify!” system [4] which has the ability to identify the important concepts in a text and then link these concepts to the corresponding Wikipedia document. However, these systems are not cross-lingual and only link important keywords to its English Wikipedia page counterpart. Many other

works or researches [5, 6, 7, 8] have also handled mono-lingual link identification from a collection of Wikipedia articles.

In the next section, the architecture of CLLD system is discussed in detail including some experimental results being shown to evaluate the performance of our CLLD system.

3 Keyword Extraction

Keyword extraction plays an important role in our Cross Lingual Link Detection system. Its goal is to retrieve all candidate words or phrases that will be linked later on with its corresponding Korean Wikipedia article.

This process, similar to that of “Wikify!” system [4], starts by generating all possible n-grams from a given input document or content. The generated n-grams could be a potential candidate keyword for the CLLD system or a nonsense phrase with no value at all. To be able to get rid of such nonsense phrases, we utilize the Wikipedia corpus (with more than 11 million English article titles including redirects as of March 2011) as our controlled vocabulary. Our controlled vocabulary will consist of all the Wikipedia article titles that are found in the corpus. This set of titles will be our basis for eliminating unwanted phrases such as “robots are” or “hiking is”, leaving us with only acceptable ones. However, this can possibly lead us to miss some potential keywords such as the morphological form of some words, e.g. “littering” or “litters” which may correspond to the article title “litter”. To avoid such restrictions, each n-grams generated are lemmatized first before eliminating those inappropriate words or phrases. To get lemmatized forms of the words, we use the WordNet API. WordNet [9] is a large lexical database of English words. Its vast list of API can be very useful depending on your needs. In our case for example, we use such tool for retrieving the lemma of a specific word. For example, for a word “littering” we can get its lemmatized form “litter”.

Another thing we need to consider is making sure that the candidate keywords we generate must have its corresponding Korean Wikipedia article. Thus, we need to filter generated keywords without a corresponding Wikipedia article.

After we had retrieved the list of all possible keywords from the previous step, a ranking scheme is then applied for each term to evaluate the value or importance of such a term in the given document. There are different types of method used to rank such keywords. One effective ranking method relies on *keyphraseness* [4] which is used in the “Wikify!” system. It takes advantage of the vast collection of Wikipedia articles which were already annotated. This method can simply be defined as “the more often a term was selected as a keyword among its total number of occurrences, the more likely it is that it will be selected again.” Keyphraseness can be estimated using the following $lp(t)$ or link probability where $cf(t)$ indicates the collection frequency of term t , and $lf(t)$ the frequency of t linked.

$$lp(t) = \frac{lf(t)}{cf(t)}$$

To rank candidate keywords, our CLLD system attempts to modify the above basic link probability. With this method, the topic dependent value of each term is also being considered aside from the original link probability score, giving certain keywords with additional ranking score. A topic dependent link probability simply means that keywords which tell more what the current input document is about will most likely be ranked higher in the set of generated candidate terms. Getting this topic dependent link probability score is done by first retrieving a co-occurrence collection of title terms and link terms. Table 1 shows a sample list of these co-occurrences. In this particular list, the keywords *Cebu* and *mango* appeared as a link term in the Wikipedia article *Philippine*. From this list, we can determine a so called 1-order score and 2-order score for our candidate keywords. A 1-order score means a direct relation of the link term to the title term. 2-order on the other hand is the indirect relatedness of such link term to a title term and is scored lower than 1-order.

Table 1. A sample list of the Link Term and Title Term co-occurrences

Link Terms	Title Terms
Cebu	Philippines
Mango	Philippines
Rajah Humabon	Magellan
Tarsier	Bohol
Magellan	Cebu
Philippines	Azkals
Bohol	Chocolate Hills

To calculate such topic dependent scores (either 1-order or 2-order), first we need to retrieve the corresponding title terms of each generated candidate keywords using the co-occurrence list, like the one shown in table 1 with link terms as the basis for matching. For example, a generated keyword list which consists of (*Cebu*, *mango*, *Rajah Humabon*, *tarsier*) will have a corresponding title terms (*Philippines*, *Magellan*, *Bohol*). These title terms will be considered as the 1-order terms. From each of these 1-order terms, we then retrieve another set of title terms (*Azkals*, *Chocolate Hills*) which will be considered as our 2-order terms. Candidate keywords will be scored accordingly using these sets, meaning candidate keywords that exist in the 1-order set of terms will be given a 1-order score and those existing in 2-order set will be given a 2-order score. Notice that the title term *Magellan* did not appear in our 2-order set. This is because a term that is in 1-order set cannot co-exist in the 2-order set, to make sure that candidate keywords will always be given the highest equivalent score.

To decide whether a candidate term t is a keyword or not, we used the following formula. $lp(t)$ is the link probability score and $co(t)$ is the co-occurrence score, which is either 1-order score or 2-order score. $\delta(p)$ is defined to 1 if p is true, 0 otherwise. θ is a threshold value. w_1 and w_2 are weight values respectively for $lp(t)$ and $co(t)$, sum of which should be 1.

$$\delta \left(\frac{w_1 lp(t) + w_1 co(t)}{w_1 + w_2} \geq \theta \right) \quad (1)$$

4 Translation Disambiguation

For each candidate terms generated by our keyword extraction process, we then link each of these terms to its corresponding Korean article. To do such thing, we utilize the whole Korean Wikipedia collection which is consisting of about 300,000 articles. From these collections we created English to Korean dictionary by retrieving all Korean article titles and its equivalent English article titles. Link to the corresponding English page of this article are usually located in the left side pane of the Wikipedia page. We used this kind of translation set for the fact that the Cross Lingual Link Detection system only link keywords to a Korean Wikipedia article, thus ensuring us that the generated translations are bound within the Wikipedia corpus.

The above Wikipedia-based translation resource may not provide translation candidates for many English keywords. To handle this, we expanded the set of English-to-Korean (E-K) translation pairs using general E-K electronic dictionaries and Korean WordNet.

One of the big challenges in natural language processing is the ambiguity of a word. For example, a word *apple* may mean “fruit” or “the Apple company”; or in other language the Korean word 사과 may mean *apple fruit* or *apology*. Such problem can be resolve by word sense disambiguation which is a way determining the correct sense of an ambiguous word.

There are several approaches used to solve WSD problems. One is a knowledge-based approach which relies on linguistic resources such as dictionaries or any set of information drawn from a sense inventory. An example method for this approach is the LESK algorithm [10] wherein the correct sense of an ambiguous word is determined by measuring the overlap of the context of the ambiguous word and the dictionary definitions or in our case, the Wikipedia articles. Another approach is the Corpus based approach which employs feature vectors (e.g. part-of-speech labels, word collocation, grammatical relationships) to be used by a machine learning classifier.

In this paper, we used a knowledge based approach in order to disambiguate ambiguous words. For our system, an ambiguous word pertains to those English words with two or more Korean translation (more than one corresponding Korean Wikipedia article). For each of these translations or Korean articles, we need to find a closest match to which the ambiguous candidate keyword exactly refers to by computing for the similarity of its context with the context of the input document. Korean document with highest similarity will be considered as the matching Korean Wikipedia document to which the candidate keyword will be linked. To compute for the similarity score for these contexts, we employ the weighted Jaccard similarity measure [11].

For example, assume that the following two sentences (A) and (B) are two contexts to be compared for WSD; (A) Apple is the pomaceous fruits of the apple tree. It is one of the most widely cultivated tree fruits. (B) I always pick apple from our

backyard most of the time. The tree bears a lot of fruits. To calculate for the weighted Jaccard similarity score of contexts A and B, first we convert both context into its vector representation with its elements having frequencies of terms occurring in A and B. Then we divide the sum of each element's minimum frequencies by the sum of their maximum frequencies which are defined as follows.

$$\text{WeightedJaccard}(A, B) = \frac{\sum_i \min(A_i, B_i)}{\sum_i \max(A_i, B_i)}$$

The disambiguation method for the CLLD system starts off by determining which word from the list of candidate keywords are to be disambiguated. These words are determined by their translation count which is in fact those candidate keywords with more than one translation or corresponding Wikipedia article. We consider these translations as the senses of the ambiguous keyword and for each of these senses we compute the similarity score of its Wikipedia article with the content of the input document using the weighted Jaccard similarity. Since the contexts of these translations are written in Korean, we first need to translate the input document to Korean by expanding each English term with its all possible Korean translation candidates which are obtained from our aforementioned E-K translation resource. The sense document or the translation with the highest similarity score will be considered as the exact translation of the ambiguous candidate keyword and thus the keyword will be linked to that corresponding Korean Wikipedia article.

5 Evaluation

For the evaluation of our E-K CLLD system, 300 English topic documents were randomly sampled from English Wikipedia dump files. These 300 documents constitute our test set. Terms linked in the topic document were considered as answer keywords for the evaluation of keyword extraction. Then, for each keyword in the 30 topic documents sampled from the test set, a human have determined its corresponding Korean Wikipedia documents. The result was used for the evaluation of translation disambiguation.

Table 2 shows the result of keyword extraction using a 3-fold cross-validation method. The proposed method based on co-occurrence scores was slightly better than the keyphraseness method, indicating a more promising approach which recognizes the effect of co-occurrence information in discriminating keywords. Room for improvements for this approach is high, thus if considered, would later yield a far better score than its current state.

Table 2. Performance of keyword extraction

Method	Precision	Recall	F-measure
Keyphraseness [2]	54.41	55.95	55.16
The proposed method	58.74	52.86	55.65

The accuracy for E-K translation disambiguation was 86.6%, showing a promising result for the feasibility of the CLLD system.

6 Conclusion

In this paper, we had presented a system that is able to automatically discover keywords within a given context and link it to a document which is written in a different language. This paper could contribute for demonstrating the use of Wikipedia as one of the resources for accomplishing different natural language processing task. Another is for providing a tool for an automatic link discovery which references each keyword to a related document with a different language.

In our future work, we will try a cross-lingual link detection system which retrieves important keywords from a given Korean context and link it to a corresponding English Wikipedia article.

References

1. Tang, L., Geva, S., Trotman, A., Xu, Y., Itakura, K.: Overview of the NTCIR-9 Crosslink Task: Cross-lingual Link Discovery. In: Proceedings of the 8th NTCIR Workshop Meeting (2011)
2. Grineva, M.: Extracting Key Terms from Noisy and Multi-theme Documents. In: WWW, Madrid, Spain (2009)
3. Milne, D., Witten, I.H.: Learning to Link Wikipedia. In: Proceedings of 17th International Conference on Information and Knowledge Management, pp. 509–518 (2008)
4. Mihalcea, R., Csomai, A.: Wikify! Linking Documents to Encyclopedic Knowledge. In: Proceedings of 16th Conference on Information and Knowledge Management, Lisboa, Portugal (2007)
5. West, R., Precup, D., Pineau, J.: Completing Wikipedia's Hyperlink Structure through Dimensionality Reduction. In: Proceedings of 18th International Conference on Information and Knowledge Management, pp. 1097–1106 (2009)
6. Kaptein, R., Serdyukov, P., Kamps, J.: Linking Wikipedia to the Web. In: Proceedings of 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 839–840 (2010)
7. Nomoto, T.: Two-Tier Similarity Model for Story Link Detection. In: Proceedings of 19th International Conference on Information and Knowledge Management, pp. 789–798 (2010)
8. West, R., Precup, D., Pineau, J.: Automatically Suggesting Topics for Augmenting Text Documents. In: Proceedings of 19th International Conference on Information and Knowledge Management, pp. 929–938 (2010)
9. Miller, G.A.: WordNet: An On-line Lexical Database. *International Journal of Lexicography* 3, 235–244 (1990)
10. Lesk, M.: Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In: SIGDOC Conference (1986)
11. Chierichetti, F., Kumar, R., Pandey, S., Vassilvitskii, S.: Finding the Jaccard Median. In: Proceedings of SODA Conference, pp. 293–311 (2010)

Efficient Finding Relationship between Individuals in a Mass Ontology Database

Dongmin Seo, Hee Kwan Koo, Seungwoo Lee, Pyung Kim,
Hanmin Jung, and Won-Kyung Sung

Korea Institute of Science and Technology Information (KISTI),
335 Gwahnagno, Yuseong-gu, Daejeon, Korea 305-806
{dmseo,hkkoo,swlee,pyung,jhm,wksung}@kisti.re.kr

Abstract. Recently, as ontology is used in the semantic web, the amount of ontology instances is increasing explosively. Today's world is complex, and so are the relationships within most knowledge domains. Especially a very complicated relationship cannot easily be overlooked. So, Experts to need decision making and information analysis expend a lot of time and effort on finding the relationships between individuals in many situations. To detect and extract relationships between individuals of interest with the help of a machine, the relationship finder system, namely RelFinder, is proposed. However, RelFinder cannot find all relationship because of its constraints for the system response time. Therefore, we propose the novel relationship finding system, named OntoRelFinder. OntoRelFinder utilizes schema paths and SPARQL endpoints to find the relationship between ontology individuals.

Keywords: Semantic Web, Finding Relationship, Ontology, RDF, RDFS.

1 Introduction

Recently, as the importance of the semantic web is being magnified, the semantic web technologies are being utilized in many different fields. The term "Semantic Web" is often used more specifically to refer to the formats and technologies that enable it. These technologies include the Resource Description Framework (RDF), a variety of data interchange formats (e.g. RDF/XML, N3, Turtle, N-Triples), and notations such as RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain [1].

RDF is a standard model for data interchange on the Web. RDF extends the linking structure of the Web using URIs that name the relationship between things as well as the two ends of the link (this is usually referred to a "triple"). This linking structure forms a directed and labeled graph, where the edges represent the named link between two data, represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations [2]. RDFS is a set of classes with certain properties using the RDF extensible knowledge representation language and providing basic elements for the description of ontologies

[3]. Therefore, we can grasp the conceptual model of RDF through RDFS. Recently, as many semantic web applications are being developing, the amount of data represented in RDF and RDFS is increasing [4].

Today's world is complex, and so are the relationships within most knowledge domains. Especially a very complicated relationship cannot easily be overlooked. Also, a large number of relationships cannot be cognitively grasped. Relationships do not follow logical human thinking. Finally, indirect relationships are hard to derive by purely cognitive reasoning [5, 6]. So, experts to need decision making and information analysis expend a lot of time and effort on finding the exact relationships between data in many situations.

In this paper, we propose a novel relationship finding system, named *OntoRelFinder* which utilizes schema path and SPARQL endpoints to find relationships between individuals. *OntoRelFinder* quickly find various and complicated relationships between two individuals and dramatically reducing unnecessary accesses to individuals.

2 The Problems of *RelFinder*

To detect and extract relationships between individuals of interest with the help of a machine, the relationship finding system, named *RelFinder*, was proposed. *RelFinder* provides a user interface to explore the huge RDF individual set by providing a means to find connection between different individuals.

The process of *RelFinder* is summarized as follows. First, selected elements are mapped to objects in the Ontology databases. Then, the objects are used as starting objects to find relationships between them. The found relationships are visualized. Finally, it can be explored by the user.

RelFinder searches for relationships between the given objects by SPARQL queries. *RelFinder* searches iteratively for connections with increasing length, starting from zero because the shortest connection is not known in advance. However, this constraint causes the degradation of search performance. In addition, because multiple changes in the direction of the relationships are difficult, *RelFinder* is only looking the simple-way relationships [5, 6]. So, *RelFinder* cannot find all relationship. This might result in wrong decisions and unwanted consequences.

3 Proposed Method

3.1 The Difference between Our Method and *RelFinder*

RDFS is built upon RDF. RDFS is a set of classes with certain properties using the RDF extensible knowledge representation language and providing basic elements for the description of ontology. The basic concept of *OntoRelFinder* is based on the fact that relationships between individuals are constrained by relationships between their classes, i.e., schema paths. So, if we utilize schema paths and SPARQL endpoints, we can quickly find various and complicated relationships between individuals.

Fig. 1 shows an example of RDF and RDFS representation of a university ontology. In Fig. 1(a), the relationship between *University* class and *Faculty* class has *Univeristy-Department-Faculty*. So, in Fig. 1(b), the relationship between *Univ#1* and *Fac#1* can only be found by the SPARQL query of Fig. 2(a). In other words, to find the relationship between *Univ#1* and *Fac#1*, only the individuals of *Department* class are considered. RelFinder, on the other hand, processes the complex SPARQL query of Fig. 2(b) because RelFinder don't know the connection between *Univ#1* and *Fac#1*. In conclusion, OntoRelFinder dramatically reduces unnecessary accesses to individuals by utilizing schema paths while finding relationships between two individuals.

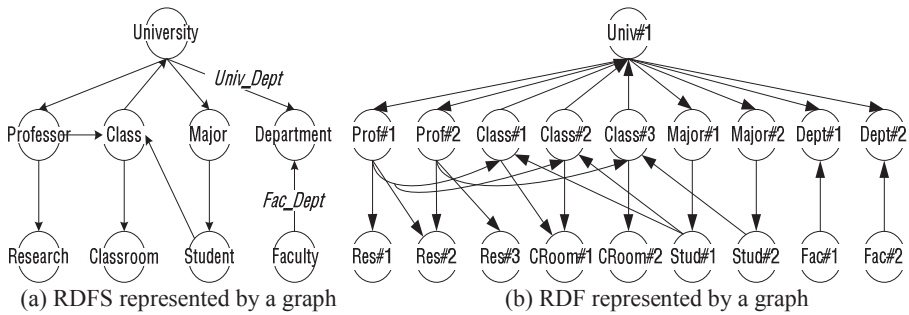


Fig. 1. An example of RDF and RDFS representation of a university ontology

```
SELECT ?d WHERE {
  Univ#1 Univ_Dept ?d.
  ?d Fac_Dept Fac#1.
}
```

(a) OntoRelFinder

```
SELECT ?d1 ?d2 ... WHERE {
  Univ#1 ?p1 ?d1. ?d1 ?p2 Fac#1. } UNION {
  Univ#1 ?p1 ?d1. ?d1 ?p2 ?d2. ?d2 ?p3 Fac#1 } UNION {
  ... }
}
```

(b) RelFinder

Fig. 2. SPARQL queries to finding the relationships between *Univ#1* and *Fac#1*

3.2 Implementation of Our Method

Fig. 3 shows the architecture of OntoRelFinder. As shown in Fig. 5, *User Interface* is implemented in Adobe Flex [7] and runs in all Web browsers. *Parameter Analyzer* receives target individuals entered by a user and gets URIs and class names of the individuals. *Graph Generator* visualizes relationships between the individuals. *CRC* finds all schema paths between two classes and stores them in *Class Relationship DB*. *CRA* finds the relationships between individuals based on the schema paths in *Class Relationship DB* and sends the results into *Graph Generator*. *RAS* gets all possible schema paths from *Class Relationship DB*. *CRAP* makes SPARQL queries using each schema path obtained from *RAS* and URIs of target individuals from *Parameter Analyzer*. The SPARQL queries are evaluated in *Triple DB*.

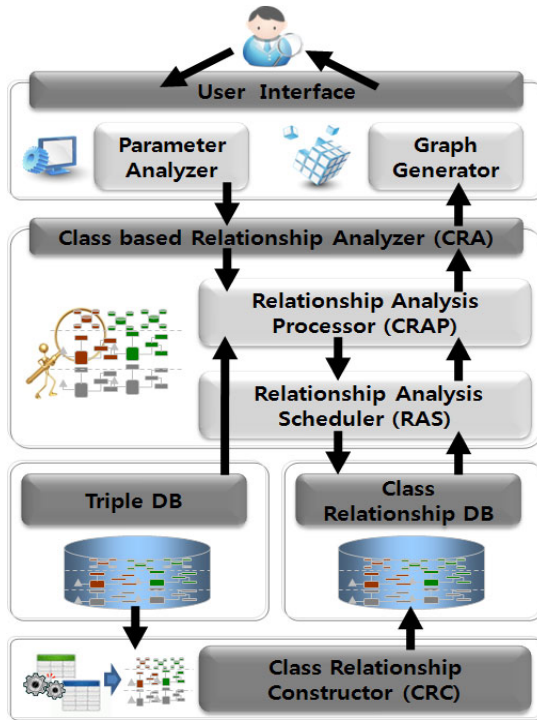


Fig. 3. The architecture of OntoRelFinder

The core of OntoRelFinder is to utilize the relationships between classes while finding relationships between two individuals. Fig. 4 shows the algorithm to find the relationships between classes in Triple DB. This algorithm uses four data structure. First, `visitedClassList` get all class names visited while processing this algorithm (Line: 01). `edgeClassList` get all class names which are located in the edge node of a graph (Line: 02). `connectClassList` get all class names which are located in the connect node of a graph (Line: 03). Firstly, this algorithm gets any class to traverse a RDFS graph. In RDFS, class is represented by `rdfs:class`. So, this algorithm gets subject of fields that has `rdf:type` as predicate of fields and `rdfs:class` as object of fields in Triple DB (Line: 07). Then, it stores the class paths from the selected class to the visited class while traversing child classes and descendant classes of the selected class by `findChildRelationship_Classes` (Line: 12). The child class and descendant class mean that these classes have the relationship with the selected class. In RDFS, the relationship between classes is represented by `rdfs:domain` and `rdfs:range`. So, this algorithm gets subject of fields that has `rdfs:domain` or `rdfs:range` as predicate as fields and the visited class as object of fields. This subject represents all relationships names related with the visited class. Then, it gets object of fields that has `rdfs:domain` or `rdfs:range` as predicate as fields and the obtained relationship names as subject of fields. This object is classes with the relationships on the visited class. This method

Algorithm 1: *findRelationship_Classes*

```

Input: Triple DB, pa_TripleDB
Output: Class Relationship DB, pa_ClassPathDB
01:  ArrayList<String> visitedClassList = new ArrayList<String>();
02:  ArrayList<String> edgeClassList = new ArrayList<String>();
03:  ArrayList<String> connectClassList = new ArrayList<String>();
    // Vertex is a data-structure for storing class
04:  Vertex rootVertex;
05:  while(visitedClassList isn't including all classes in pa_TripleDB)
06:    rootVertex = new Vertex();
07:    rootVertex = get the a class in pa_TripleDB. It isn't included in visitedClassList;
08:    if(rootVertex == NULL)
09:      break
10:    End if
11:    visitedClassList.add(rootVertex);
12:    findChildRelationship_Classes(pa_TripleDB, pa_ClassPathDB, rootVertex);
13:    inferCycleRelationship_Classes(pa_PathDB, rootVertex);
14:    rootVertex = NULL;
15:  End while
16:  findEdgeRelationship_Classes(pa_PathDB, edgeClassList, connectClassList);

```

Fig. 4. The algorithm to find the relationships between classes on *OntoRelFinder*

ignores the direction of the relationship between classes because this method has no meaning among the directions of *rdfs:domain* and *rdfs:range*. So, *OntoRelFinder* can find the complex relationships between classes easily.

4 Conclusion

This paper proposes *OntoRelFinder* to efficiently detect and extract relationships between individuals of interest in the huge ontology database. *OntoRelFinder* utilizes schema paths and SPARQL endpoints to find the relationship between individuals. So, *OntoRelFinder* dramatically reduces unnecessary accesses to individuals while finding relationships between two individuals. Also, we describe the performance of *OntoRelFinder* is better than that of *RelFinder* through their methodologies. Finally, *OntoRelFinder* is applied to *Relationship Path* service of InSciTe [8] that is developed by KISTI. *Relationship path* discovers and visualizes associated relationships among technologies and agents to support technology opportunity discovery.

In the future, we will compare OntoRelFinder with the existing system in various experiments to verify the superiority of OntoRelFinde. Also, we plan to apply our method on many different fields.

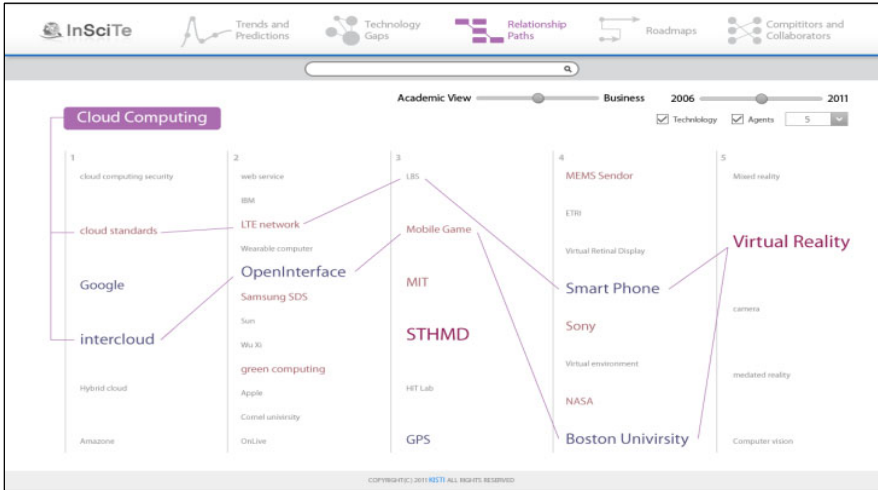


Fig. 5. Relationship paths among technologies and agents in InSciTe service

References

1. Herman, I.: W3C Semantic Web Activity. W3C (2008), <http://www.w3.org/2001/sw/>
2. Brian, M., Hewlett, P.L.: RDF Primer. W3C Recommendation (2004), <http://www.w3.org/TR/rdfsyntax/>
3. Brian, M., Hewlett, P.L.: RDF Vocabulary Description Language 1.0: RDF Schema. W3C Recommendation (2004), <http://www.w3.org/TR/rdf-schema/>
4. Sandro, H., Ivan, H., Eric, P.: W3C Semantic Web Activity. W3C Recommendation (2008), <http://www.w3.org/2001/sw/>
5. Philipp, H., Sebastian, H., Jens, L., Steffen, L., Timo, S.: RelFinder: Revealing Relationships in RDF Knowledge Bases. In: Chua, T.-S., Kompatsiaris, Y., Mérialdo, B., Haas, W., Thallinger, G., Bailer, W. (eds.) SAMT 2009. LNCS, vol. 5887, pp. 182–187. Springer, Heidelberg (2009)
6. Steffen, L., Philipp, H., Timo, S., Jiirgen, Z.: The RelFinder User Interface: Interactive Exploration of Relationships between Objects of Interest. In: 15th International Conference on Intelligent User Interfaces, pp. 421–422 (2010)
7. Adobe Systems Incorporated (2009), <http://www.adobe.com/products/flex>
8. Lee, S., Lee, M., Jung, H., Kim, P., Kim, T., Seo, D., Sung, W.-K.: Supporting Technical Decision-Making with InSciTe. In: International Semantic Web Conference (2010)

Generating Knowledge Map for Acronym-Expansion Recognition

Do-Heon Jeong, Myung-Gwon Hwang, and Won-Kyung Sung

Department of Information Technology Research
Korea Institute of Science and Technology Information(KISTI)
305-806 Daejeon, Korea
{heon,mgh,wksung}@kisti.re.kr

Abstract. In this paper, we present a method for instance mapping and URI resolving to merge two heterogeneous resources and construct a new semantic network in a viewpoint of Acronym-Expansion. Acronym-Expansion information extracted from two unstructured large dataset can be remapped by using linkage information between instances and measuring string similarity. All the semantic resources and systems are utilized for constructing infrastructures of KISTI's TOD, a new information analytics project based on semantic technologies.

Keywords: Instance mapping, URI resolution, Acronym-Expansion, Knowledge map, Technology opportunity discovery(TOD).

1 Introduction

Semantic Web techniques has been developed constantly and the techniques could be attained at a degree to be applied to various fields such as analyzing business opportunities like emerging technologies in the near future [1,2,3]. With regard to the trends, Korea Institute of Science and Technology Information (KISTI) is carrying out a TOD (Technology Opportunity Discovery) research project which is one of the activities for information analytics based on the semantic technologies [3]. The TOD project is a kind of serial infra-structure including system, data, analysis methodology to support R&D environment for researcher groups and enterprises. For this huge project, roles of the semantic techniques and its application technologies are very important; one is about extracting core parts from a large amount of knowledge information and amalgamating with the refined data, and the other is about analyzing the data for intelligent and semantic services. Above all in the semantic techniques and its application areas, one of the core and fundamental fields is a linguistic processing as ever.

This paper contains a method to manage knowledge resource which is capable to be applied to extracting and analyzing technical terms. We explain the method with examples of whole process about knowledge management such as constructing acronym-expansion information, clustering similar entities, and assigning URIs to the entities. We strongly expect that this smart knowledge is adapted usefully for fundamental systems extracting fine-grained information.

2 Motivations

2.1 Technical Term Extraction (Information Extraction)

Information extraction techniques are very important to analyze the large scale database. Therefore, the NLP processing is on the lowest level of TOD's architecture. KISTI's previous work related to NLP proposed a domain independent terminology recognition system based on machine learning method using dictionary, syntactic features, and Web search results all together [4]. Since term instances, however, obtained from the system do not contain semantic information like word senses, additional methods are necessary as well such as WSD (Word Sense Disambiguation) techniques (highly related to instance alignment and URI resolution), and authority control to term variants.

2.2 AcroDic: Acronym Dictionary System

In order to extract terms from the huge information, knowledge base of significant terminologies is essential for information analysis. In previous work, to support more precise natural language processing for unstructured documents, AcroDic has been developed. The method consists of three parts: first part extracting pairs of acronym and its expansion, second part reliable selecting acronym-expansion pairs, and the last part connecting appropriate glossaries of the selected pairs [5,6]. This paper contains a part of the method about the construction of the AcroDic which employs two large data resources.

2.3 Instance Mapping and URI Resolving

There are many ontology mapping techniques and research has been conducted actively in this area. These techniques range from simple string comparisons to complex machine learning approaches [7]. Instance mapping ontology enables to transform data from different representations into the target ontology with the complex mapping relations. Because ontology mapping is an error prone and time consuming task, automatic mechanisms to reduce the human effort is required in this field [8]. In this paper, though we do not handle instance mapping techniques in a viewpoint of the ontology scheme directly, but our knowledge base system manages all kinds of representations about general lexical resources as well as Acronym at upper level.

3 Generating Knowledge Map for Acronym-Expansion

3.1 Data Collection and Construction Method

Data resources (Acronym-Expansion) in this research are constructed from two types of domain. One is the product of collective intelligence, Wikipedia¹ and the other is

¹ <http://en.wikipedia.org/>, <http://dbpedia.org/>

scholarly database for KISTI's NDSL information service². Especially, NDSL stores massive metadata up to 50 million papers and we have obtained initial Acronym-Expansion dataset by pattern recognition methods from about 13 million abstracts. Methods for automatic acquisition of the initial dataset are as follows:

Table 1. Date Resources and Extraction Methods

	Wikipedia	NDSL database
Feature	1. Wikipedia as knowledge base 2. Acquisition of expansions for acronyms via disambiguation pages 3. Acquisition of synonyms of expansions via URI redirection page	1. Scholarly paper for NDSL 2. Acquisition of many term variations from free text with ease. 3. Acquisition of commonly occurring errors and omissions
Result Data Sets	Acronym : 33,101 Meaning Group : 135,449 Expansion : 512,587	Acronym : 30,672 Meaning Group : 75,890 Expansion : 89,409
Extraction Method	[step1] extracting acronym-expansion pairs from disambiguation page [step2] extending expansions to synonyms (remapped Wiki labels with Wiki URI)	[step1] extracting acronym-expansion pairs from free text [step2] clustering term variants (assigned URI to cluster automatically)
Clustering Synonyms	Expanding synonyms by Wiki semantic linkage information	Measuring N-gram similarity betw. expansions to control term variations (Dice's coefficient ≥ 0.9)
Ex) Term Variants & Synonyms	x-ray diffraction, x-ray diffuse scattering, x-ray scattering, resonant anomalous x-ray scattering	x-ray diffraction, x ray-diffraction, x-ray dffraction, x-ray diffaction, x-ray differaction, x-ray diffllaction, ...

3.2 Instance Mapping Methods

In this part, we explain some methods to handle acronyms and expansions from two large domains. These cases are very simple and easy to understand, but at all stages, synchronization techniques to keep or change the assigned past URI and property for instance management need to be considered carefully.

Case 1. No String Match, No Mapping

In case that there is no string matching or similarity value does not exceed thresholds between any instances(both of the acronym and the expansion), instances' URIs do not need to be changed and have only to move them onto the integrated network. There are no URI changes in the system.

Case 2. The Same Acronyms, No Mapping between Expansions

In case that more than two acronyms have the same string value and no expansions for each of the acronyms have the same string value, we can assign the same URI

² <http://www.ndsl.kr/>

to the acronym only. At this time, other expansions have to keep the past URI. At all stages of instance mapping and URI resolving, interconnected triple data have to be updated to the new URI information.

Case 3. The Same Acronyms, Mapping between Expansions

The same acronyms which have the same string value are given the same URI, and the same expansions with exactly matched string or high similarity value are only given the same URI.

Case 4. Different Acronyms Sharing more than One Expansion

When different acronyms share one expansion at least, the past URIs on acronyms are kept and shared expansions are resolved as the same URI.

Case 5. Complex Network considering Variations

The examples of ‘Case 1 to 4’ occur frequently in the mapping process. But in the real data environment, we may encounter more complicated situations like Fig.1. Node A1:E2 (means Expansion 2 linked to Acronym 1) and A2:E2 can accompany many variants. As we mentioned, In Wikipedia, term variants can be gathered from Wiki URI pages because many synonymous terms redirect to the same Wiki URI (but it may not be tight related terms). In NDSL, terms’ variation problem can be solved because singular-plural nouns, punctuations (like hyphen, space, etc.), mistyping (like errors and omissions), and so on occur frequently and abundantly in the documents. The variant helps to tie the same meaning group through the string comparison between extended nodes as shown in Fig. 1(b).

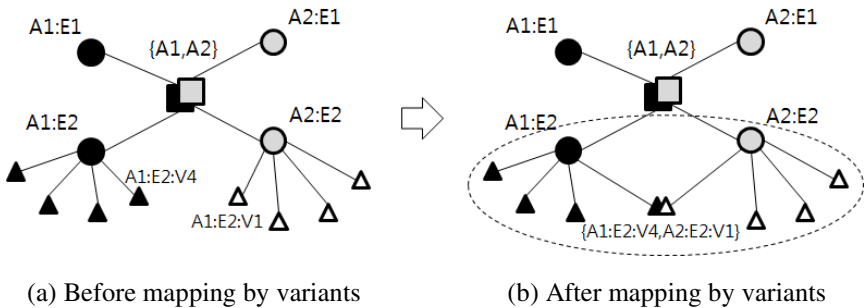


Fig. 1. Grouping the same meaning group by extended term variants
 (■ : Acronym, ● : Expansion, ▲ : Variation)

3.3 Marking the Meaning Group on the Semantic Network

This is the method to discriminate the meaning groups on the semantic network. We inserted specific marks (which is a kind of dummy node) to the link between an acronym and a group of expansions and its variants (see dot-circled area in Fig. 2).

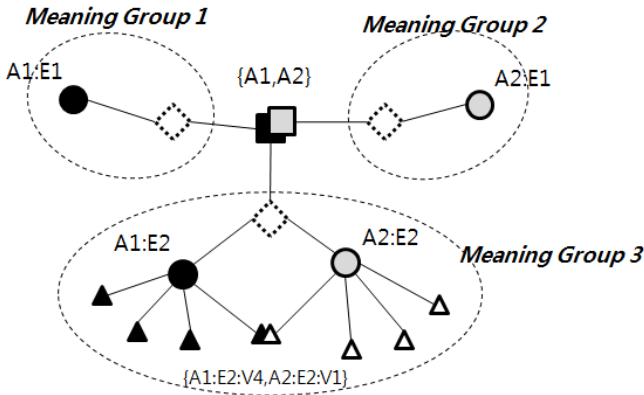


Fig. 2. Knowledge map using meaning group marks (=dummy nodes for bundle URI)

After disambiguating senses, one bundle URI is assigned to one meaning group in addition. Because meaning group marks help to depict the structure of semantic network well, human and computer are able to understand the topology of the semantic relationships with ease.

3.4 In the Real Semantic Network

The knowledge map which describes the real world, as shown in Fig. 3, contains the complicated relationship between instances on the semantic network. Every instance in the same meaning group (MG) has its own instance URI and each of the MGs has the same bundle URI. In the Fig. 3, total 5 meaning groups exist on the map as below:

MG #1 : {XRD@1}

MG #2 : {XRD@2, XRDS@1}

MG #3 : {XRD@3, XRDS@2, XDS@1, RASX@1, XRS@3}

MG #4 : {XRS@1}

MG #5 : {XRS@2}

This knowledge base system not only presents expansions for the acronym, but returns the preferred expansion for the acronym. Query starting with the acronym “XRD” ends up with the final expansion “X-ray scattering” as the following process,

Step 1. Recognize acronym with NLP tools, then input query acronym : “XRD”

Step 2. Retrieve all candidates (Meaning Groups) : “{XRD@1:{X-ray diffractometer, ...}, XRD@2:{X-ray diods, ...}, XRD@3:{X-ray diffraction, X-ray scattering, ...}}”

Step 3. Select one group from candidates by Clue Words : “XRD@3:{X-ray diffraction, X-ray scattering, ...}”

Step 4. Return the final preferred term (based on term frequency etc.) : “X-ray scattering”

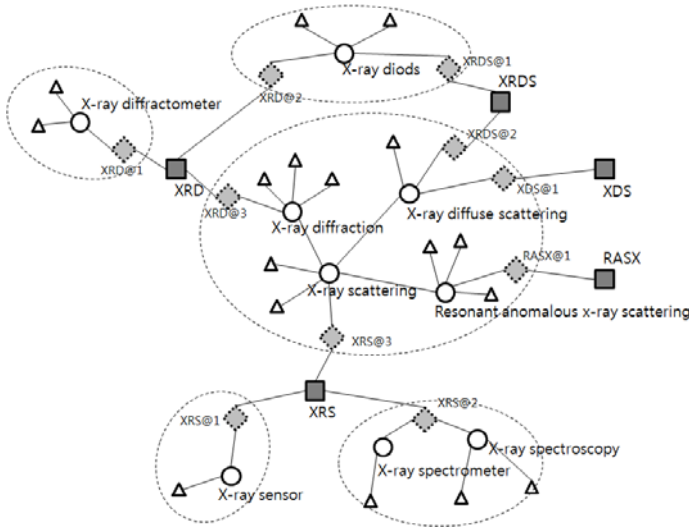


Fig. 3. Semantic network for Acronym-Expansion (sense disambiguated)

By instance mapping process, acronym set is merged about 18.6% and meaning groups are only mapped about 2.4%, but we assume that language resources of a variety of expansions and variants are built up in each of the meaning groups (see Table 2).

Table 2. Results of URI resolution

	WIKI	NDSL	WIKI+NDSL	Mapping Ratio
Acronyms	33,101	30,672	51,884	18.6% (11,889)
Meaning Groups	135,449	75,890	206,196	2.4% (5,143)
Expansions	512,587	89,409	596,405	0.9% (5,591)

4 Conclusions

In this paper, we have proposed the knowledge map generating method to support and enhance information extraction techniques for the information analysis. Based on the large-scaled two heterogeneous knowledge bases (Wikipedia and NDSL), we have constructed very large language resources and knowledge map which are consisted of Acronym-Expansion, and discriminated the meaning of Acronym on the semantic network. This system can contain general lexical resources as well.

In the future work, we will prove that the word sense disambiguation techniques improve the performance of information analysis based on text mining tasks.

References

1. CUBIST: Your Business Intelligence, Project home page at, <http://www.cubist-project.eu/> (viewed August 22, 2011)
2. IARPA: Be The Future, FUSE project home page at, http://www.iarpa.gov/solicitations_fuse.html (viewed August 22, 2011)
3. Lee, M.K., Jung, H.M., Kim, P., Sung, W.K.: Technology Intelligence Service for Establishing R&D Strategy. *Journal of Korea Institute of Information Scientists and Engineers* 17(5), 337–341 (2011)
4. Choi, Y.S., Song, S.K., Chun, H.W., Jeong, C.H., Choi, S.P.: Terminology Recognition System based on Machine Learning for Scientific Document Analysis. *The Transactions of Korea Information Processing Society* (Submitted)
5. Hwang, M.G., Jeong, D.H., Sung, W.K.: Automatic Construction of Acronym Knowledge Base and its Application Methods. *The Journal of Korean Institute of Information Technology* 9(6), 167–176 (2011)
6. Hwang, M.G., Jeong, D.H., Sung, W.K.: AcroDic 1.0: Acronym Dictionary. In: *Proceedings of The First Conference On Terminology, Language, and Content Resources*, pp. 39–48 (2011)
7. Shamdasani, J., Bloodsworth, P., Munir, K., Rahmouni, H.B., McClatchery, R.: MedMatch-Towards Domain Specific Semantic Matching. In: *19th International Conference on Conceptual Structures (ICCS)*, pp. 375–382 (2011)
8. Mocan, A., Cimpian, E.: Mapping Creation Using a View Based Approach. In: *1st International Workshop on Mediation in Semantic Web Services*, pp. 97–112 (2005)

Inference Verification for Enhancing Confidence in Semantic Information Service

Seungwoo Lee, Mikyoung Lee, Hanmin Jung, and Won-Kyung Sung

Korea Institute of Science and Technology Information,
245 Daehangno, Yuseong-gu, Daejeon, 305-806, Korea
{swlee, jerryis, jhm, wksung}@kisti.re.kr

Abstract. Recent semantic information services based on ontology provide users with very useful analytic information but do not often give intuitive understanding to users. This may lead to decrease users' confidence in semantic information services. To solve this issue, we suggest inference verification method which explains how an implicit fact was obtained. Inference verification is defined as a set of facts and rules caused to derive an implicit fact. Practicality and visualization are considered as important factors to implement this task since the task is for real users. We are convinced that the suggest method can give confidence in semantic information services to users and plan to evaluate it through user test.

Keywords: Inference Verification, User Confidence Enhancement, Triple Dependency Indexing, Visualization.

1 Introduction

Recent information service is gradually evolving into information analytics service, jumping over the limit of traditional information search. This change is enabled by semantic technologies such as text mining, ontology, inference and semantic repository. For example, recent semantic information services such as OntoFrame [1], iLaw [3], RKBexplorer [4] and RelFinder [5] can discover and provide useful analytic information and implicit knowledge – e.g., domain experts, researchers' network, and relational path between entities – from explicit RDF (Resource Description Framework) [11] data using semantic technologies. These semantic information services are very useful to users since they can help users to save time and labor in analyzing large amount of data. However, users often have doubts about the confidence in the semantic information services because users cannot get any explanation for how such implicit and analytic information was obtained from the services, while users could verify directly the relevance of results from traditional information search services. In other words, a non-explainable semantic information service may cause to decrease users' satisfaction at the service.

To have users get confidence, semantic information service could explain the whole processes of semantic information analysis to users. In particular, since an ontology-based information analytic service is usually implemented with a SPARQL [12] query and pos-processing of the query results, one possible solution is to present

to users a description of the service and ontology data matched to the SPARQL query. This corresponds to the proof lane at the third-top layer in the Semantic Web layer cake¹ as shown in Fig. 1. The proof means verification of (a) trustability of ground facts and (b) correctness of inference processes. (a) requires evaluating authority of source of the ground facts and (b) requires checking whether inference processes follow formal logic and approved rules or not. Focusing on the latter, this paper presents an inference verification method as a partial proof to enhance users' confidence in semantic information service. Inference verification is defined as a set of facts and rules which caused to derive an implicit fact and explains how such an implicit fact was obtained when a semantic information service contains inferred facts. Practicality and effective visualization are considered as important factors to implement this task since the task is for real users.

The remaining parts of this paper are organized as follows. The next section presents some related works with comparison to ours. Section 3 and section 4 explains practical method and visualization for inference verification respectively. Finally, conclusion is given in section 5.

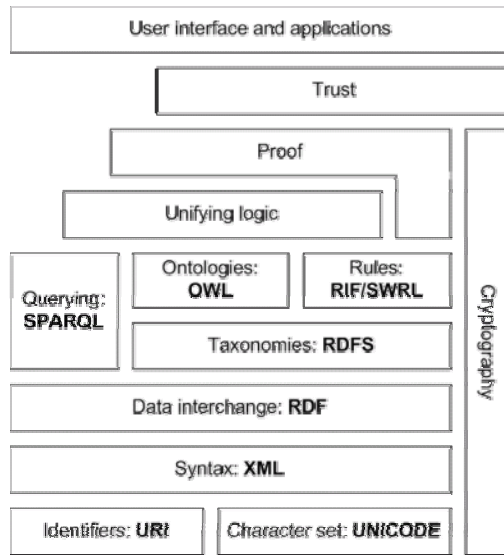


Fig. 1. Semantic Web layer cake

2 Related Work

Previous researches related to inference verification have been studied mainly with focus on the aspect of ontology engineering. Parsia et al. [6] presented a method for detecting and diagnosing errors in ontologies, especially, T-Box. They detect unsatisfiable concepts inferred by Pellet, a DL reasoner, and diagnose unsatisfiability

¹ Semantic Web Layer Cake, http://en.wikipedia.org/wiki/Semantic_Web_Layer_Cake.

and inconsistency of ontology. However, they did not consider the scalability of ontology, especially large-scaled A-Box. Kalyanpur et al. [7] and Suntisrivaraporn et al. [8] also introduced a method for debugging inconsistencies or undesired subsumption in OWL DL [12] ontologies. Both of them defined justification as a set of ground facts causing OWL entailments and suggested DL-based methods that find all justifications of doubtful OWL entailments for their explanations. In addition, Suntisrivaraporn et al. considered efficiency and scalability of the proposed method using locality-based modules. Horridge et al. [9] presented a tool for finding justification in the form of a workbench plug-in for Protégé 4. As a kind of ontology modeling tool, it checks whether the ontology was modeled consistently or not and explains the reason, if some inconsistencies occurred, to support ontology debugging.

These all works follows DL-based inference and finds justifications on demand (on-line) for debugging OWL DL ontologies. That is, when some errors like unsatisfiable concepts and inconsistencies in OWL DL ontologies were detected, they provide the ground facts caused the errors to make it easy for users to fix them. However, justification is often not sufficient for ontology engineers to understand the concerned OWL entailment when it passes through several intermediate processes of inference since the intermediate processes are not explained in justification. To solve this issue, a justification-oriented proof method was suggested [10]. This method also focuses on the aspect of ontology modeling and aims to help ontology engineers to debug OWL DL ontologies with explanation of intermediate steps of OWL entailments. This approach is appropriate for ontology engineers who are familiar to DL and OWL, but it is not easy for general users who have no such prior knowledge to understand the processes of OWL entailments without clear explanation of applied inference rules.

In contrast to the previous work, our method focuses on the aspect of confidence in semantic information services, not ontology modeling, and suggests inference verification approach which supports fast retrieval and visualization of ground facts and inference rules with intermediate steps caused to derive the concerned facts. For practicality, we follow rule-based inference, although it cannot support all OWL DL entailments, and suggest triple dependency indexing and searching method. In addition, we visualize ground facts and inference rules with intermediate steps in graphical form for general users with no prior knowledge about OWL to understand the meaning.

3 Triple Dependency Indexing and Searching

In this section, we explain the triple dependency indexing and searching method for practical inference verification. For the sake of convenient explanation, we use an imaginary ontology and inference rules which model academic research field as shown in Fig. 2 and Fig. 3, respectively. In particular, rule-based forward inference approach is adopted since it is much more practical than DL-based inference, especially in large-scaled ontologies. The relationships such as *'hasDomain'* and *'expertIn'* and class membership of *'QualifiedArticle'* can be inferred by the rules in Fig. 3. Each rule has a description template for use in visualization.

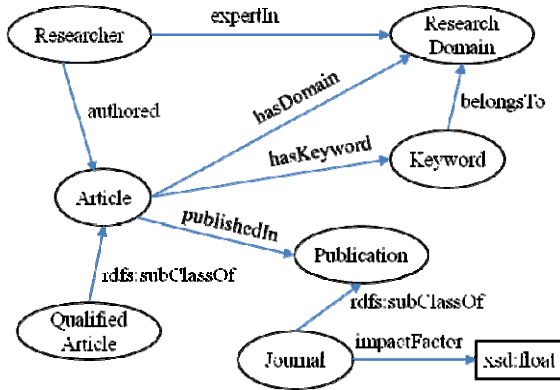


Fig. 2. Schema of an imaginary ontology

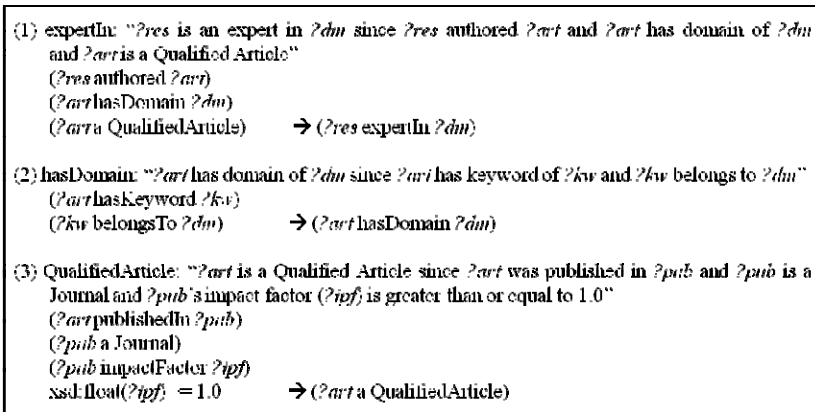


Fig. 3. Inference rules

In an ontology, a fact is represented as a triple consisted of subject, predicate and object. Rule-based forward inference is a process that derives new facts by applying inference rules to the existing facts. Therefore, there exist dependencies by an inference rule between a new fact and existing facts. Fig. 4(a) shows an example of ontology instances following the schema in Fig. 2 and Fig. 4(b) shows new facts inferred by applying the rules to the instances. The serial numbers at the left of each fact and each rule indicate fact ID’s and rule ID’s respectively. The inferred facts depend on the instances in Fig. 4(a) and inference rules in Fig. 3. For instance, fact (22) depends on fact (6), fact (16), fact (17) and rule (3), i.e., according to rule (3), “(22) *Art1 is a Qualified Article since (6) Art1 was published in Pub1, (16) which is a Journal and (17) whose impact factor (3.0) is greater than 1.0*”. The inferred facts can also be used for further inferences. For example, fact (20) and fact (22) are

inferred ones and fact (24) depends on these two inferred facts again. A new fact can be inferred from different existing facts and rules. For instance, fact (20) can be inferred by applying rule (2) to fact (4) and fact (12) or fact (5) and fact (14).

These inference dependencies can be indexed to present inference verification efficiently during inference. Like document indexing, where each document is indexed with its keywords, we index fact dependencies with inferred facts. To implement this triple dependency indexing method, we first assign unique IDs to each fact and rule, and then use the IDs, when indexing each inferred fact, to save space and time taken to store and retrieve the triple dependencies. Fig. 5 shows an example of triple dependency indexing corresponding to the ontology instances in Fig. 4. The dependency can be indexed in two ways: forward and backward. Backward indexing is sufficient for inference verification but forward indexing may also be useful for supporting incremental inference, which is not explained here since it is out of the scope of this paper.

Using the triple dependency index, we can fast and correctly retrieve facts and rules caused to infer the implicit fact when receiving a request on verification of an implicit fact. The retrieving process is applied recursively until no more retrieved facts are inferred ones.

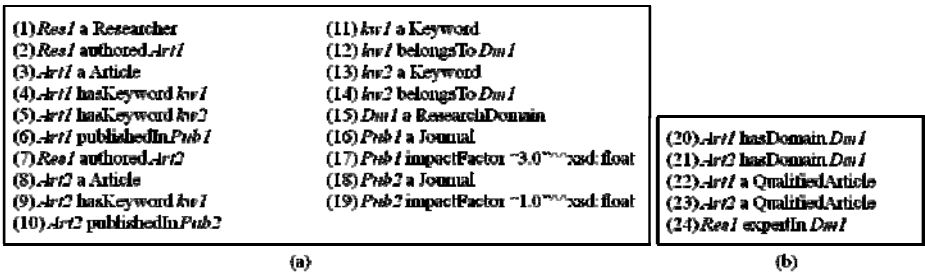


Fig. 4. Ontology instances: (a) ground facts, (b) inferred facts

InferredTripleID	RuleID	TripleID1	TripleID2	TripleID3	TripleID	InferredTripleID	TripleID	InferredTripleID
20	2	4	12	-	4	20	17	22
20	2	5	14	-	12	20	2	24
21	2	9	12	-	5	20	20	24
22	3	6	16	17	14	20	22	24
23	3	10	18	19	9	21	7	24
24	1	2	20	22	12	21	21	24
24	1	7	21	23	6	22	23	24
					16	22		

Fig. 5. Triple dependency indexing: (a) backward index², (b) forward index

² For simplicity of explanation, we assume that a triple is inferred by maximum three depending triples.

4 Visualization

This section presents effective visualization of inference verification. Visualization should be intuitive enough for general users to understand since general users may have no prior knowledge about ontology and inference. So, we adopt graphical form with textual description as shown in Fig. 6, which explains the visualization of verification on an inferred fact “*Res1 expertIn Dm1*” by stages.

Fig. 6(a) is the start stage of the visualization and the relationship ‘*expertIn*’ represented with dotted arrow indicates that it was obtained by inference and can be verified by clicking ‘⊕’. If that is clicked, the verification on the relationship is visualized as in Fig. 6(b). That is, facts supporting the relationship are added to the visualization and the rule is explained with a tooltip text, which is obtained from the description template of the rule. Then, even general users can understand why *Res1* is an expert in *Dm1*. This verification still contains another four inferred relationships, two ‘*hasDomain*’ and two ‘*a*’. These four relationships can be verified further if users want. Fig. 6(c) shows a stage of visualization after clicking both relationships of ‘*hasDomain*’ and ‘*a*’ of ‘*Art1*’.

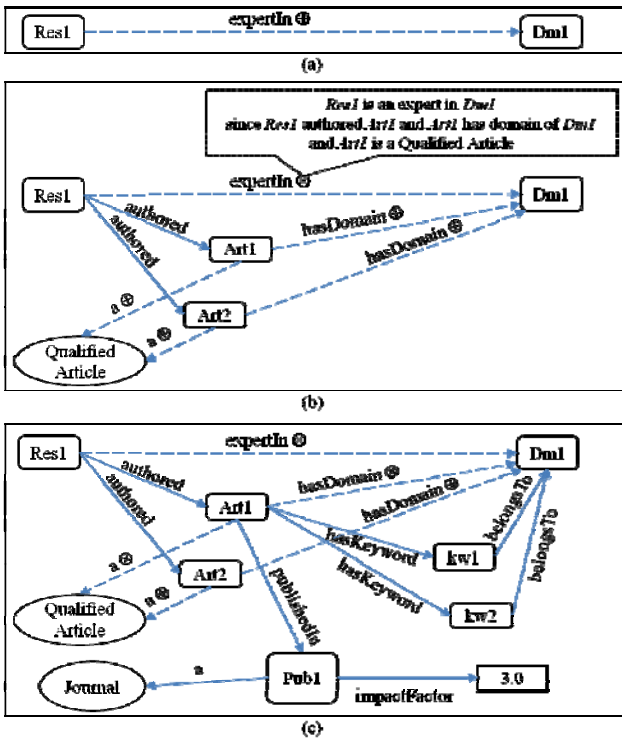


Fig. 6. Visualization of inference verification on “*Res1 expertIn Dm1*”: (a) initial, (b) after clicking ‘*expertIn*’ relationship, (c) after clicking both relationships of ‘*hasDomain*’ and ‘*a*’ of ‘*Art1*’

5 Conclusion

Recent semantic information services based on ontology provide users with very useful analytic information but do not often give intuitive understanding to users. This may lead to decrease users' satisfaction with semantic information services. To solve this issue, we suggested inference verification which indexes triple dependencies including rules, when explanation of an implicit fact is requested, searches the triple dependencies and visualizes the processes of inference to effectively explain how the implicit fact was inferred. We are convinced that this will give confidence in semantic information services to users and plan to evaluate it through user test.

References

1. Use Case: OntoFrame 2008— A Semantic Portal Service of Academic Research Information (2008), <http://www.w3.org/2001/sw/sweo/public/UseCases/OntoFrame/>
2. Jung, H., Lee, M., Kim, P., Sung, W.-K.: Technology-Based Decision-Making Support System. In: Salvendy, G., Smith, M.J. (eds.) HCII 2011, Part II. LNCS, vol. 6772, pp. 262–267. Springer, Heidelberg (2011)
3. Jung, H., Kim, P., Lee, S., Lee, M., Seo, D., Sung, W.-K.: iLaw system: Intelligent Legislation Support System based on Semantic Web and Text Mining Technologies. In: 5th International Conference on Methodologies, Technologies and Tools enabling e-Government (MeTTeG 2011), Italy, pp. 1–12 (2011)
4. Glaser, H., Millard, I.C., Jaffri, A.: RKBExplorer.com: A Knowledge Driven Infrastructure for Linked Data Providers. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 797–801. Springer, Heidelberg (2008)
5. Philipp, H., Sebastian, H., Jens, L., Steffen, L., Timo, S.: RelFinder: Revealing Relationships in RDF Knowledge Bases. In: Chua, T.-S., Kompatsiaris, Y., Mérialdo, B., Haas, W., Thallinger, G., Bailer, W. (eds.) SAMT 2009. LNCS, vol. 5887, pp. 182–187. Springer, Heidelberg (2009)
6. Parsia, B., Sirin, E., Kalyanpur, A.: Debugging OWL Ontologies. In: Ellis, A., Hagino, T. (eds.) 14th International Conference on World Wide Web (WWW 2005), pp. 633–640. ACM, Chiba (2005)
7. Kalyanpur, A., Parsia, B., Horridge, M., Sirin, E.: Finding All Justifications of OWL DL Entailments. In: Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L.J.B., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., Cudré-Mauroux, P. (eds.) ASWC 2007 and ISWC 2007. LNCS, vol. 4825, pp. 267–280. Springer, Heidelberg (2007)
8. Suntisrivaraporn, B., Qi, G., Ji, Q., Haase, P.: A Modularization-Based Approach to Finding All Justifications for OWL DL Entailments. In: Domingue, J., Anutariya, C. (eds.) ASWC 2008. LNCS, vol. 5367, pp. 1–15. Springer, Heidelberg (2008)
9. Horridge, M., Parsia, B., Sattler, U., Schneider, T.: Working with Explanations of OWL Entailments. In: Tutorial of 8th International Semantic Web Conference (ISWC 2009), Washington, DC (2009)
10. Horridge, M., Parsia, B., Sattler, U.: Justification Oriented Proofs in OWL. In: Patel-Schneider, P.F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J.Z., Horrocks, I., Glimm, B. (eds.) ISWC 2010, Part I. LNCS, vol. 6496, pp. 354–369. Springer, Heidelberg (2010)
11. Resource Description Framework (RDF), <http://www.w3.org/RDF/>
12. SPARQL Protocol and RDF Query Language (SPARQL), <http://www.w3.org/TR/rdfsparql-query/>
13. Web Ontology Language (OWL), <http://www.w3.org/TR/owl-features/>

Design of TOD Model for Information Analysis and Future Prediction

Jinhyung Kim, Seungwoo Lee, Jinhee Lee, Mikyoung Lee, and Hanmin Jung

52-11 Eoeun dong, Yuseong gu, Daejeon, Republic of Korea
{jinhyung,swlee,jhlee29,jerryis,jhm}@kisti.re.kr

Abstract. Analyzing mass information and supporting insight based on analysis results are very important work but it needs much effort and time. Information analysis and future prediction about science and IT filed data are also very critical tasks for researchers, government officers, businessman, etc. Therefore, in this paper, we propose technology opportunity discovery (TOD) model based on feature selection and decision making for effective, systematic, and objective information analysis and future forecasting of science and IT field.

Keywords: Technology Opportunity Discovery, Information Analysis, Information Prediction, Feature Selection, Decision Making.

1 Introduction

The precise information analysis and new opportunity discovery are very important for future prediction, future countermeasures decision, and future plan establishment. However, as the amount of information in science and IT field increases exponentially every year, data analysis about that information or extraction new opportunity from documents, papers, patents, etc. becomes more difficult and complicate. Until now, there have been several researches regarding information analysis of mass data and new opportunity discovery [1-3]. Traditional studies focused on information analysis and conclusion deduction based on the scenario method, the Delphi method or AHP method. Above methods are based on non-systematic process and depend on subjective opinions of experts. The scenario method [4,5] is a strategic planning method where a group of experts analyze base information for decision making. However, because there are several irregular and arbitrary cases in the scenario processing, reliability is very low. The Delph method [6,7] is a structured communication technique, originally developed as an interactive forecasting method which relies on a panel of experts. In the Delphi method, experts answer questionnaires in two or more rounds. After each round, a facilitator provides an anonymous summary of the experts' forecasts from the previous round as well as the reasons they provided for their judgments. However, because the Delphi method depends on subjective opinions from experts, that cannot guarantee credibility of prediction results. The AHP method [8,9] is a structured technique for organizing and analyzing complex decision based on mathematics and psychology. Like the Delphi method, the AHP method depends on subjective view of

an evaluator or group of experts. As a result, the AHP method also cannot assure accuracy of forecasting results.

To overcome the limitations mentioned above, many systematic and objective methods are suggested such as Foresight and Understanding form Scientific Exposition (FUSE) [10], Combining and Uniting Business Intelligence with Semantic Technology (CUBIST) [11], Text Mining Software for Technology Management (VantagePoint) [12], and so on. These projects aim to support decision making by analysis and pattern recognition of scientific documents. However, many researches and projects focus on information analysis which is insufficient to support new opportunity discovery or future prediction.

Therefore, Korea Institute of Science and Technology Information (KISTI) [13] have been researched regarding information analysis on science and technology field, and technology opportunity discovery since 2010. The research is named InSciTe and, information analysis and technology opportunity prediction are based on TOD Model. We will describe TOD Model at the next section in detail. In this paper, we analyze research trends based on papers and patents and execute feature selection. Using several features extracted by feature selection, we can recognize technology trends and predict future prospective technologies.

This paper is organized as follows. In the section 2, we describe TOD Model architecture. Section 3 represents snapshot of InSciTe service including TOD Model. Section 4 concludes this paper with future works.

2 Technology Opportunity Discovery Model Architecture

Technology discovery model consists of 3 sub-models: technology life cycle discovery (TLCD) model, technology development speed discovery (TDSD) model, and emerging technology discovery (ETD) model. The TLCD model decides emerging phase of a specific technology through feature selection and analysis extracted from papers and patents information. We adopt emerging phase in the Gartner's hype cycle. Emerging phase is comprised of 5 steps: technology trigger, peak of inflated expectations, through of disillusionment, slope of enlightenment, plateau of productivity. First, The TLCD model extracts several features from papers and patents. Then the TLCD model performs feature selection and construct decision tree for emerging phase decision. The decision tree can be optimized by machine learning and we can finally decide emerging phase of technologies. The TDSD model decides technology development speed of a specific technology. In other words, the TDSD model calculates time (year) for moving to final 5th step. The TDSD model uses the exponential moving average (EMA) method [14] for calculating technology development speed. EMA is a variant of moving average that is, more weight is given to the latest data. The exponential moving average is also known as exponentially weighted moving average. The EMA method is mainly utilized for analysis and prediction of stock price. By technology speed, the TDSD model can predict when the technology can arrive at the final plateau step. That is, if period to the final plateau step of a specific technology is short than others, that the technology can be decided as an emerging technology. The ETD model decides emerging technology among lots of technologies in science and IT field. To decide emerging technologies, we define

following 5 types of emerging technology (ET): Trigger Emerging Technology, Associated Emerging Technology, Pre-stable Emerging Technology, Referenced Emerging Technology, and Derived Emerging Technology. Each emerging technology is defined as follows.

- Trigger ET is technologies new appeared in 2 years and have bigger growth rate than average growth rate.
- Associated ET is technologies associated and similar technologies of top technology in 2 years & technologies which have direct relationship by ontology inference.
- Pre-stable ET is technologies which arrive at the final plateau step in 2 years.
- Reference ET is technologies referenced commonly from other emerging technologies.
- Derived ET is technologies appeared newly and derived by emerging technology such as augment reality derived from mobile device / in-memory DB management derived from mobile device.

The ETD model decides emerging technology based on above 5 types of emerging technology.

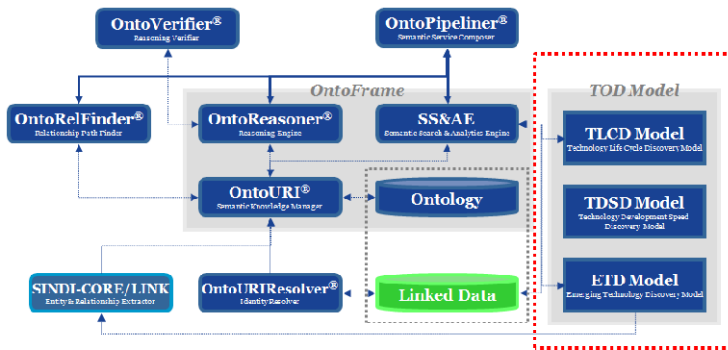


Fig. 1. Architecture of TOD Model

3 Service Snapshot

As we described in introduction, TOD model is applied to Intelligent Science and Technology (InSciTe) service developed by KISTI since 2010. Fig. 2 shows main page of InSciTe service. As shown in main page, several technologies are represented on the hype cycle graph. We calculate feature values about all of technologies (96,000 technologies). Then, the ETD model decides emerging technologies among 96,000 technologies based on 5 types of emerging technology defined as shown in section 2. Determined emerging technologies are represented at the main page of InSciTe service as Fig. 2. The TLC model decides step of emerging technologies based on decision tree. All of emerging technologies can be classified into one of 5 emerging steps: Technology trigger, peak of inflated expectation, though of disillusionment,

slope of enlightenment, plateau of productivity. Finally, the TDSD model calculates development speed of emerging technologies based on EMA method. The development speed of emerging technologies decided by the TDSD model is represented as 4 steps such as '>', '>>', '>>>', '>>>>'. The more '>' symbols means the faster development speed of technology.



Fig. 2. Main Page Snapshot of InSciTe Service

Except for the main page, if a user selects a specific technology, InSciTe service provides 8-years (5-years for the past, 3-years for the future) emerging steps of the specific technology. This result is supported by the TLCD model.



Fig. 3. Snapshot of InSciTe Service

4 Conclusion

In this paper, we defined TOD model for effective information analysis and future prediction based on papers and patents information. Compared to conventional method, projects, and services, the suggested TOD model supports more systematic process and objective analysis/prediction results. The TOD model consists of TLCD model, TSD model, and ETD model for much more diverse information analysis and prediction information provision to interested users.

As future works, we will perform simulation test with diverse datasets. Gartner's hype cycle is released in many research areas and many other research center and governments predict emerging technology lists. Simulation test with several kinds of dataset will improve prediction accuracy much higher.

References

1. Richard, S.: Patent Trends as a technological Forecasting Tool. *World Patent Information* 5(3), 137–143 (1983)
2. Dereli, T., Durmusoglu, A.: A trend-based patent alert system for technology watch. *Journal of Scientific & Industrial Research* 68(8), 674–679 (2009)
3. Users and abuses of patent statistics. In: Pavitt, K. (ed.) *Handbook of Quantitative Studies of Science and Technology*. Elsevier (1998)
4. Wright, G., Goodwin, P.: Decision making and planning under low levels of predictability: Enhancing the scenario method. *International Journal of Forecasting* 25(4), 813–825 (2009)
5. Hetmanska, K., Nguyen, A.: A method for learning scenario determination and modification in intelligent tutoring systems. *International Journal of Applied Mathematics and Computer Science* 21(1), 69–82 (2011)
6. Okoli, C., Pawlowski, S.: The Delphi method as a research tool: an example, design considerations and applications. *Information and Management* 42(1), 15–29 (2004)
7. Hanafizadeh, P., Mirzazadeh, M.: Visualizing market segmentation using self-organizing maps and fuzzy Delphi method- ADSL market of a telecommunication company. *Expert Systems with Applications* 38(1), 198–205 (2011)
8. Liu, G., Jin, Y., Li, F.: The application of AHP method in well control risk evaluation by controllable factor analysis. *Journal of Southwest Petroleum University* 33(2), 137–141 (2011)
9. Duran, O.: Computer-aided maintenance management systems selection based on a fuzzy AHP approach. *Advances in Engineering Software* 42(10), 821–829 (2011)
10. Foresight and Understanding for Scientific Exposition (FUSE), http://www.iarpa.gov/solicitations_fuse.html
11. Combining and Uniting Business Intelligence with Semantic Technology (CUBIST), <http://www.cubist-project.eu>
12. VantagePoint, <http://www.thevantagepoint.com>
13. Korea Institute of Science and Technology Information (KISTI), <http://www.kisti.re.kr>
14. Song, F., Hao, S., Hao, M.: Study on NC instruction interpretations algorithm with highorder differentiability based on moving average. *Mechatronics and Intelligent Materials* 900-903 (2011)

A the Effects of SCM Activities on Manufacturing Capabilities and Manufacturing Performances

Su-Hyun Jung^{*}, Sung-Hong Kim, and Sung-Lyong Kang

College of Commerce & Business, Chungbuk National University, 410 Sungbong-ro,
Heungduk-gu, Cheongju, Chungbuk, 361-763 Republic of Korea
{jungsuhyun, kimsh, slkang}@chungbuk.ac.kr

Abstract. Recently, dramatic changes and increasing uncertainties in business environment have raised awareness of SCM(supply chain management) for efficient resource management between individual companies and their suppliers. SCM has been actively studied so far, and many investigators have made efforts to establish major activities or success factors involved in supply chains. However, few studies have considered the relevance of SCM activities to manufacturing capabilities and corporate performance. As a rule, SCM is about an integrated control of overall value chains involving a company, its suppliers, distributors and customers. The present study, however, is to empirically analyze the effects of SCM activities on manufacturing capabilities and supply chain collaboration performance by narrowing down the focus of consideration to inter-relations between individual companies and their suppliers. That is, major SCM activities taking place between individual companies and their suppliers are divided into separate and shared efforts, so as to establish their effects on manufacturing capabilities and corporate performance.

Keywords: SCM, Activities, Manufacturing Capabilities, manufacturing Performance.

1 Introduction

Today, companies confront cut-throat competition in global market environment and focus on how to provide their products for customers in a timely manner and at a lower price due to shorter product life cycles and higher customer expectations. To survive fierce competition, companies try to build a close collaborative system by linking manufacturers, suppliers and customers together in a chain. In the past, manufacturers used to take the initiative in one-sided relations with their suppliers, but since the advent of global market, they have had to make the best quality of goods possible at the lowest cost possible to brace for global competition, which inevitably has required competition between supply chains, not between individual companies. Subsequently, to hold competitive advantages against other supply chains, an

^{*} First Author: Ph.D. Candidate, Graduate School of Business (phone: 82-10-4411-9057);

increasing number of companies are turning to SCM (supply chain management) that links their external suppliers with customers.

SCM has been actively investigated so far, and lots of investigators have tried to establish major activities or success factors in relation to supply chains. Still, the relevance of major activities of supply chains to manufacturing capabilities and corporate performance has been hardly studied. Accordingly, the present study finds out major activities required to improve SCM performance, and considers causal relevance of such activities to manufacturing capabilities and supply chain collaboration performance in domestic manufacturers adopting SCM.

To begin with, the present study performs a literature review on major SCM activities, manufacturing capabilities and corporate performance, and develops hypotheses and a study model to delve into the relevance of SCM activities derived from the literature review to manufacturing capabilities and corporate performance. To test the hypotheses, a corporate questionnaire survey was done targeting the personnel in charge of SCM or purchasing in manufacturing companies. Data collected were analyzed in terms of validity and reliability and the hypotheses were tested.

2 Theoretical Background

2.1 SCM Structure

SCM (Supply Chain Management) is applied to a range of fields including products, services, design, production and delivery, and it has a network structure consisting of raw material suppliers, producers, warehouses, distribution centers, wholesalers, retailers and customers.

A supply chain is a network involving all entities affecting its performance such as suppliers of a supplier and customers of a customer and manages relations between members to maximize profitability of every entity in the chain. Figure 1 shows a general structure of a supply chain.

2.2 SCM Activities

SCM-related activities and success factors have been well documented by many investigators. Some previous studies on SCM-related major activities dealt with causal relationships between success factors and business management performance, while others suggested critical factors for SCM performance using information-based approaches [1, 6, 9, 11, 17, 18, 22, 23, 24, 25, 27].

Among overseas studies, Anderson et al. (1997) concluded that 7 capability factors determined a successful SCM such as integrated supply chain operation, efficient logistics networks, efficient logistics network application, supply chain flexibility, strategic relationship with suppliers, performance measurement and information

system implementation[1]. Peter Gilmour (1999) asserted that top management’s customer-oriented support for supply chains, efficient logistics management, demand-oriented sales planning, lean production, supplier partnership, integrated SCM, integrated information system, improved technology, integrated performance measurement, team operation and appropriate organizational structure influenced SCM performance via supplier-buyer relationships[25]. Table 1 outlines the previous findings on SCM activities.

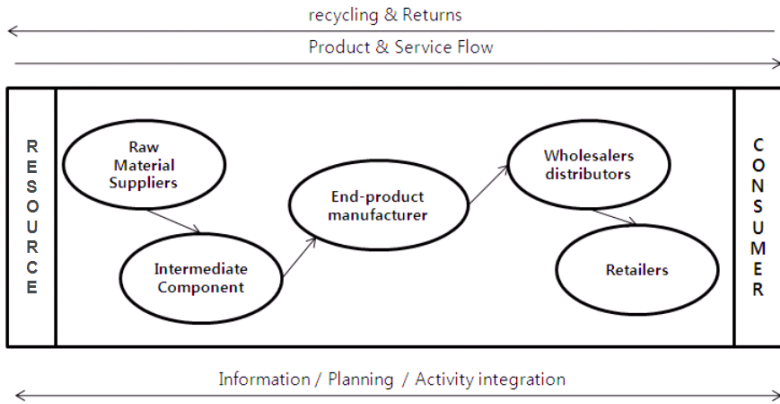


Fig. 1. General structure of a supply chain

Source: Jeol D. Wisner., Keah-Choon Tan., G. Keong Leong(2008) Principles of Supply Chain Management, McGraw-Hill. p.7[15]

Table 1. Previous findings on SCM activities

Investigator Factor		1	2	3	4	5	6	7	8	9	10
Separate activities	Top management’s support				0	0	0				0
	Human resource development				0			0	0		
	Logistics management		0		0						
Shared efforts	Partnership					0		0	0	0	0
	Performance sharing	0						0	0		0
	Common goal management	0		0	0		0	0		0	

Investigators : 1. Ellram(1995), 2. Anderson(1997) 3. Metz(1997), 4. Peter Gilmour(1999), 5. Marien(2001), 6. Goo J. Y. (2001), 7. Kim Ch. B. (2002), 8. Kim J.S. (2004), 9. Gah H. G. (2004), 10. Park Y. W. (2004).

As presented in Table 1, SCM activities are comprised of not only individual efforts to hold each company’s competitive edge through inter-corporate collaboration and exchange but also shared efforts to create win-win synergies. This study includes

top management’s support, human resource development and logistics management in the individual efforts, and partnership, performance compensation and common goal management in the shared efforts.

2.3 Manufacturing Capabilities

A number of studies have investigated SCM-related activities and manufacturing capabilities. Hayes and Wheelwright (1984) suggested basic manufacturing capabilities such as quality of labor force, management skills, product quality, employee participation, manufacturing engineering and continuous improvement as necessary elements required to build superior manufacturing capabilities[12]. Anderson et al.(1997) used logistics cost reduction, shorter order lead time, inventory reduction, sales growth and equipment effectiveness improvement as part of management performance[1]. Yoon H. D. et al. (2001) measured management performance in terms of the levels of customer service improvement, cost saving, operation process improvement and sales improvement. To measure these levels, they used product quality, delivery and flexibility as measurement variables[30]. Park Y. W. et al. (2004) used a BSC (balanced score card) model in their study on SCM performance measurement and success factors and included measurement items such as product quality, flexible response and on-time delivery[24]. Kim S. M. et al. (2005) measured corporate performance in line with different types of integrated SCM in light of delivery lead time, product development capability and product diversification[15]. Lee J. H. et al. (2007) studied supply chains by categorizing the relevance of product quality, delivery pace, process flexibility and price leadership as part of combinative capability[20].

It is necessary to see these manufacturing capabilities from the existing perspectives of products, delivery, flexibility and price in order to find out causal relations between major SCM activities and manufacturing capabilities. Hence, this study attempted to measure manufacturing capabilities via such factors as product quality, delivery pace, process flexibility and price leadership based on previous studies. Table 2 outlines these elements vis-a-vis previous findings.

Table 2. Previous studies on manufacturing capabilities

Investigators Factors		Hayes & Wheelwright (1984)	Anderson (1997)	Yoon H. D. et al. (2001)	Park Y. w. et al. (2004)	Kim S. M. et al. (2005)	Lee J. H. et al. (2007)
Manufacturing capabilities	Manufacturing quality	0		0	0	0	0
	Delivery pace		0	0	0	0	0
	Process flexibility			0	0	0	0
	Price leadership		0			0	0

2.4 Performance

Diverse studies have been performed on corporate performance following supply chain activities. In a study on 5 companies driving SCM, Elliff(1998) applied the performance of production activities and set cost saving, quality improvement, shorter cycle time, higher efficiency of production equipment, shorter cycle for new product development, higher efficiency in inventory and equipment allocation and better customer service as part of management performance[5]. Schlegel(1999), in his study on CSF and management performance, adopted increase in profitability, improvement in customer service and added values, equipment efficiency improvement, inventory reduction, cycle time shortening and cost saving as variables for performance measurement[26]. Chen and Antony (2004) analyzed over 400 relevant theses and developed 4 conceptual structures in SCM. As for performance, supplier operating performance (volume flexibility, schedule flexibility, on-time delivery, reliability and consistency in delivery, quality and cost), buyer operating performance(volume flexibility, delivery pace, reliability, conformance to product specification, cost, fast confirmation of orders, customer complaint processing and customer satisfaction) and buyer financial performance(ROI, profit margin, net income before tax and present corporate value) were measured[3].

Meanwhile, Yoon H. D. et al. (2001) used inventory cost saving levels, inventory turnover improvement levels, inventory layout efficiency improvement levels, equipment efficiency improvement levels and asset efficiency improvement levels, and as for variables to measure sales improvement levels, they measured market share expansion levels, material profitability improvement levels, cash flow improvement levels and capital investment efficiency improvement levels[30]. Park Y. W. et al. (2004) applied a BSC model and used profitability improvement, cash flow improvement, sales growth and asset turnover improvement for financial performance[24]. Lastly, in a study on types of supply chain integration in connection with diversification strategies, Kim S. W. (2005) suggested cost reduction status and differentiation for management performance. Purchase and operating costs, inventory, warehouse, sales, distribution and transport were presented as the factors for cost reduction status, and on-time delivery of raw materials by suppliers, rates of returned materials, order processing pace, time for processing requests for material return and product/process innovation as the elements for differentiation[16].

Table 3. Previous studies on corporate performance

Investigators Factors		Elliff (1998)	Schlegel (1999)	Chen et al. (2004)	Yoon H. D. et al. (2001)	Park Y. W. et al. (2004)	Kim S. W. (2005)
Corporate performance	Non-financial performance	0	0	0	0		0
	Financial performance	0	0	0	0	0	0

As previous studies excluded a range of performance measures in reference to SCM collaboration performance, this study is to adopt non-financial performance measures from the perspective of product flow and financial performance measures from the perspective of profitability. Table 3 outlines these performance factors in contrast to previous findings.

As mentioned above, variables and measures selected for this study based on previous findings are as in Table 4.

Table 4. Variables and Measures

Variables	Measures
Top management's support	Top management's support for SCM activities
	Top management's contribution to SCM activation
	Top management's drive for SCM application
Human resource development	Training on SCM
	Training on concept and importance of SCM
	Performing practical training on SCM
	Training for SCM
Logistics management	Using delivery warehouse system
	Operating inventory visibility management
	Operating demand estimation management
	Adjusting product flow
Partnership	Managing partnerships with major suppliers and subcontractors
	Mutual contracts with major suppliers and sub-contractors
	Major suppliers and subcontractors exchange opinions
	Major suppliers and subcontractors collaborate
	Relationships with major suppliers and subcontractors
Performance compensation	Monetary compensation for goals achieved(e.g. incentives)
	Non-monetary compensation for goals achieved(e.g. performance rating)
	Fair compensation for goals achieved
Common goals	Seamless relationships with suppliers
	Trade relations with suppliers
	Business communication with suppliers
Product quality	Suitable quality
	Product durability
	Product reliability
	Product performance quality
	Overall quality

Table 4. (continued)

Delivery pace	Quick processing of customer order
	Just-in-time delivery of products
	Shorter time spent from order to procurement
Process flexibility	Quick change in product mix
	Quick change in production volume
	Diverse product mix using the same equipment
Price leadership	Manufacturing products at lower cost than competitors
	Price competitiveness against competitors
	Price control capability superior to competitors
Non-financial performance	Inventory reduction
	Improved inventory turnover
	Improved equipment/asset efficiency
Financial performance	Increased sales
	Increased net income for the year
	Improved cash flow

3 Model and Hypotheses

3.1 Model Design

Based on the theoretical background reviewed earlier, this section deals with the variables of supply chain collaboration, manufacturing capabilities and major activities of SCM expected to influence manufacturing capabilities and supply chain collaboration, and describes a study model and hypotheses. The study model involving the variables constructed here is in Figure 2.

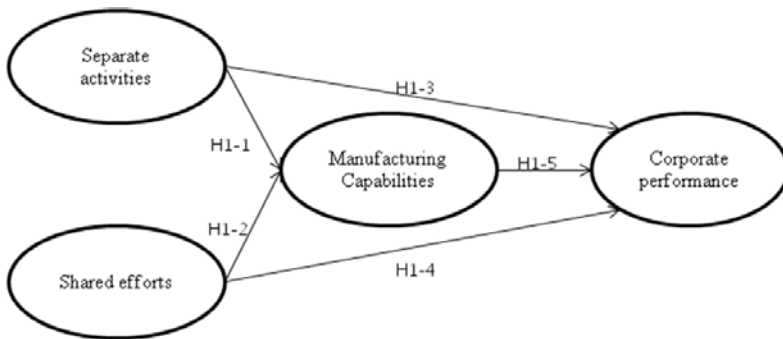


Fig. 2. The study model

4 Conclusion and Limitation

This study constructed a structural equation model to analyze the influence of SCM activities on manufacturing capabilities, supply chain performance and corporate performance. The present study is characterized by dividing supply chain activities into separate and shared activities, and by developing a model covering manufacturing capabilities to corporate performance. The measures used in this study were derived based on a literature review and modified and supplemented with the help of practitioners. Then, a questionnaire survey was carried out targeting the personnel in charge of purchase and SCM in domestic manufacturing companies. 101 responses excluding insincere ones were used in the analysis with the PLS statistical package. The analysis found the following results.

First, to raise corporate performance in individual companies, separate SCM efforts including top management's support, human resource management and logistics management activities were found significant for manufacturing capabilities and corporate financial and non-financial performance. This indicates that top management's support for SCM activities and efficiency in human resource and logistics management are important.

Second, shared SCM activities including partnership, performance compensation and common goals management were found significant for manufacturing capabilities and corporate financial and non-financial performance. This finding suggests that business entities should develop common goals among external partners, maintain suitable communication among partners and practice reasonable compensation for performance to motivate them to engage in SCM activities.

Third, as proved by previous studies on manufacturing capabilities and corporate performance, corporate capabilities were found to influence corporate performance. Therefore, by maintaining corporate activities to improve product quality, delivery pace, process flexibility and price leadership, companies can obtain efficient competitive capabilities of low cost, spot delivery, high quality and flexibility and develop corporate competitive capabilities as well as secure competitive advantages within an industry.

This study attempted an empirical analysis to verify the relevance of SCM activities to manufacturing capabilities and corporate performance.

As the questionnaire survey was conducted targeting only the respondents working in the manufacturing companies, and as the inter-corporate relations were considered without taking into account overall aspects of SCM, generalization of the population may be limited. Thus, future studies need reinforce the survey responses with a view to raising the possibility of further generalization.

References

1. Anderson, J.C., Narus, J.A.: A Model of Distributor firm and Manufacturer firm working partnership. *Journal of Marketing* 54, 42–58 (1990)
2. Bhattacharjee, A., Sanford, C.: Influence processes for information technology acceptance: an elaboration likelihood model. *MIS Quarterly* 30(4), 805–825 (2006)

3. Chen, I.J., Paulraj, A.: Towards theory of supply chain management; the constructs and measurements. *Journal of Operations Management* 22, 119–150 (2004)
4. Chin, W.W.: The partial least squares approach for structural equation modeling. In: Marcoulides, G.A. (ed.) *Mordern Method for Business Research*, pp. 295–336. Lawrence Erlbaum Associates (1998)
5. Elliff, S.A.: Organizing for Excellence: Five Case Studies. *Supply Chain Management Review*, 38–45 (Winter 1998)
6. Ellram, L.M.: Partnering Pitfalls and Success Factors. *International Journal of Purchasing & Materials Management* 31(2), 36–44 (1995)
7. Fornell, C., Bookstein, F.: Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research* 19, 440–452 (1982)
8. Fornell, C., Larcker, D.: Structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1), 39–50 (1981)
9. Gah, H.G.: A Study on the Influence of Partnership Factors on SCM Performance – Focusing on Domestic SMEs, Master’s thesis, Postgraduate School, Chungang Univ. (2005)
10. Gefen, D., Straub, D., Boudreau, M.: Structural Equation Modeling Techniques and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems* 7(7), 1–78 (2000)
11. Goo, J.Y.: SCM-based Logistics Innovation. *The Marketing Management Study* 5(3), 105–120 (2001)
12. Hayes, R.H., Wheelwright, S.C.: *Restoring Our Competitive Edge: Competing Through Manufacturing*. Wiley, New York (1984)
13. Shin, H., Collier, D.A., Wilson, D.D.: Supply Management orientation and Supplier/Buyer Performance. *Journal of Operations Management* 18, 317–333 (2000)
14. Wisner, J.D., Tan, K.-C., Keong Leong, G.: *Principles of Supply Chain Management*, p. 7. McGraw-Hill (2008)
15. Kim, S.M., Gang, S.R.: An Exploratory Study on the Influence of SCM Types and Corporate Characteristics on Performance. *The Journal of Korean Management Society* 18(4), 1553–1573 (2005)
16. Kim, S.W.: Types of Integrated Supply Chains and Linkage to Diversification Strategies. *The Study of Business Administration* 34(2), 471–496 (2005)
17. Kim, J.S., Kim, D.J.: An Exploratory Study on Developing SCM Performance Indicators for SMEs. *Korea Intelligent Information System Society, Information System Study* 13(2), 195–224 (2004)
18. Kim, C.B.: A Case Study on Successful Partnership Factors in Companies Adopting SCM. *Internet eCommerce Study* 2(2), 99–118 (2002)
19. Lee, H.L., Billington, C.A.: Managing Supply chain inventory: Pitfalls & Opportunities. *Sloan Management Review* 33(3), 65–74 (1992)
20. Lee, J.H., Ryu, C.H.: A Study on Manufacturers’ Environment, Strategy and Performance. *The Study of Business Administration* 37(3), 477–513 (2007)
21. Nunnally, J.C., Vernstein, I.J.: *Psychometric Theory*. McGraw-Hill, New York (1994)
22. Marien Edward, J.: The Four supply Chain Enablers. *Supply Chain Management Review* 4(2), 60–68 (2000)
23. Metz, P.J.: Demystifying Supply Chain Management: Accomplishments and Challenges. In: *Council of Logistics Management Annual Conference Proceedings*, pp. 50–55 (1997)
24. Park, Y.W., Lee, J.H.: A Study on SCM Performance Measurement and Success Factors. In: *Book of Theses Presented in the Winter Seminar Hosted by Korean Distribution Society*, pp. 123–152 (2004)

25. Gilmour, P.: Benchmarking Supply Chain Operations. *International Journal of Physical Distribution & Logistics Management* 29(4), 259–266 (1999)
26. Schlegel, G.L.: Supply Chain Optimization: A Practitioner's Perspectives. *Supply Chain Management Review*, 50–57 (Winter 1999)
27. Seo, A.Y., Shin, G.S.: An Empirical Study on SCM Performance Factors in View of Supplier-Buyer Relationships. *Korea Society of Management Information Systems, Information Systems Review* 3(1), 191–203 (2001)
28. Srite, M., Karahanna, E.: The Role of Espoused National Cultural Values in Technology Acceptance. *MIS Quarterly* 30(3), 679–704 (2006)
29. Tan, K.C., Kannan, V.J., Handfield, R.B.: Supply Chain Management: Supplier performance and firm performance. *International Journal of Purchasing and Materials Management* 34(3) (1998)
30. Yoon, H.D., Han, H.Y.: An Empirical Study on SCM Performance in Domestic PC Parts Suppliers. *The Journal of Korean Logistics Society* 11(1), 117–141 (2001)

A Digital Convergence Maturity Model: The Relative Importance of Factors^{*}

Seung-Jun Yeon^{**} and Sung-Hyun Hwang

Electronics and Telecommunications Research Institute (ETRI)
218 Gajeongno, Yuseong-gu, Daejeon, 305-700 Republic of Korea
sjyeon@etri.re.kr, hyun052@etri.re.kr

Abstract. This study was conducted with the larger goal of designing a maturity model for digital convergence services, needed to set a policy for their promotion. Its primary objective is to determine the relative importance of maturity indicators related to digital convergence readiness, the level of use of, and the level of satisfaction with, services and performance in four major fields in which digital convergence services are currently used: education, healthcare, transportation and public administration. The data for this study were obtained by surveying a panel of 80 experts from the four fields, and the experts were sampled based on a list, using both the purposive sampling and snowball sampling methods. Analysis was performed using the AHP technique. The results show that in education, healthcare and public administration, convergence readiness is the most important factor, followed by the level of use and level of satisfaction and performance, in this order. Meanwhile, in the field of transportation, the level of use and satisfaction proved the most important, followed by convergence readiness and performance.

Keywords: Convergence, Maturity, Readiness, Usage, Performance, AHP.

1 Introduction

The convergence between IT and other industries is taking place currently at multiple levels including between technologies, between services as well as between companies. Amid high hopes that digital convergence services will help spawn new industries, serving as the new sources of competitiveness, they are also expected to become blue ocean areas for products and services with virtually unlimited potential for value creation. Digital convergence services are already commercialized, and various government policy undertakings related to broadcasting and telecommunications are currently being implemented in line with the changing media demand at a national and societal level and the evolving policy environment. However, on the research front, there is an extreme dearth of existing literature related to strategies for promot-

^{*} This paper is a revised and extended version of a paper was presented at the 16th Annual International Conference on Industrial Engineering – Theory, Applications and Practice. This research is supported by Korean Communications Commission.

^{**} Corresponding author. Electronics and Telecommunications Research Institute, 138 Gajeongno, Yuseong-gu, Daejeon, 305-700 Republic of Korea (phone: 82-42-860-6437; fax: 82-42-860-6504;)

ing digital convergence services or evaluation of related performance benefits. More recently, there have been discussions about methods to measure the penetration of digital convergence services in different sectors of society and assess performance benefits resulting from their use, so that related information may be used for making policies for their promotion. Such methods would provide useful estimates for developing service promotion policies, and assist service providers in developing successful business models, which may be used as references for related export projects. They would also help identify new promising business areas

This study, a pre-study conducted in view of developing policy suggestions for promoting digital convergence services, attempts to develop an early framework for assessing the level of maturity among digital convergence services and determine the relative importance of measurement indicators. The framework proposed in this study is developed around three indicators - readiness, the level of use and level of satisfaction and performance - and their relative importance is measured in four major sectors in which digital convergence services are used most actively; namely, education, healthcare, transportation and public administration.

2 Designing a Converged Digital Media Service Framework

Before proceeding to develop a framework for digital convergence services, one must first define precisely what digital convergence services are. The Korea Communications Commission defined them as “new-concept services, born out of the application of convergence technology elements from the broadcasting and telecommunications sectors to existing service fields to create high market value,” while mentioning multimedia-based services and mobile services as two main components of these services. Thanks to their multimedia capabilities and mobility, digital convergence services are able to improve productivity and efficiency across all sectors and fields of our society, including education, healthcare, public administration, transportation. Various pilot programs are currently underway to test them.

Of the many fields in which digital convergence services may be productively used, these four fields have been selected by the Korean government as the priority areas for promoting their utilization. Education, public administration, transportation

Table 1. Service Fields

Field	Description
Education	Diffusion of IPTV in schools; creation of a distribution system for high-quality educational content; development of converged digital media-based education services through promotion of video conference-based education programs among others.
Public administration	Discovering new demand for multimedia-aided public services, including government-to-citizen services, local or tourist information, legal advice services.
Transportation	Creation of a new transportation service market which can promote easy access to transportation information and the smart evolution of transportation means.
Healthcare	Creation of a multimedia-based healthcare service market providing services such as disease prevention, telemedicine and healthcare information services.

and healthcare are four of the most important public interest fields, delivering essential services to Koreans, which are crucial for improving their quality of life.

Before developing a framework for these four fields, we first analyzed existing performance evaluation models and indices. Existing indices can be classified into three types: comparative index, scale-type index and maturity index. With a comparative index, a base year is chosen and is set to 100, as is the case with indices like the KOSPI and the KOSDAQ, to calculate an index value for years that follow the base year. In the case of the KOSPI index, for example, a value of 100 is assigned to the base date of January 4, 1980. The value of the composite stock price index at this reference date is set to 100, and the total market capitalization, calculated by weighting the price per share for each stock by the total number of listed shares on a given date, is compared with that of the reference date. These indices allow the analysis of price growth or decline trends by quarter or by year and can be maintained consistently and continuously.

As for scale-type indices, they are oftentimes used for calculating the national competitiveness of countries. As a general rule, indices comparing national competitiveness such as the WCS or NCR are calculated by standardizing various qualitative and quantitative indicators and applying a weight to them.

Indices of this type, although they offer the advantage of being able to take into consideration a large number of items, also have a major drawback. The evaluation items can vary depending on the period evaluated, as they are importantly affected by changes in the external environment; there is, therefore, no real guarantee of consistency with indices of this type.

Table 2. Types of ICT-related Indices

Scale-type indices	[UN] e-Government readiness index, [UN] Online participation index, [ITU] ICT development index (IDI), [ITU] Digital opportunity index (DOI), [WEF] Network readiness index, [WEF] National competitiveness index (technology readiness), [IMD] National competitiveness index (technology infrastructure), [EIU] Digital economy index, [EIU] IT industry competitiveness index
Maturity indices	[UN] Web Measurement Index

Finally, with maturity indices, maturity phases are first established for the object measured for maturity, then the current level of maturity¹ is measured so that directions to be taken to reach an ideal phase can be suggested.

Of the above-described indices, those that are used frequently in the ICT fields are listed in <Table 2> above:

As has been already said, methods for evaluating performance, although varied, all have their respective innate limitations. In this study, we will present a digital convergence maturity model (DCMM), developed by combining several methods, in a

¹ The phase considered the most ideal in the current point-in-time was assigned 100.

manner to maximize the respective advantages of each of them and remedy their associated disadvantages.

This model was developed under three principal criteria: The first is that the index was designed drawing on existing evaluation indices. When evaluating existing indices, we paid particular attention to their relevance to digital convergence services. The second criterion was that the design of the maturity model should enable the identification of future improvement opportunities. The third and last criterion was the requirement of considering performance variables.

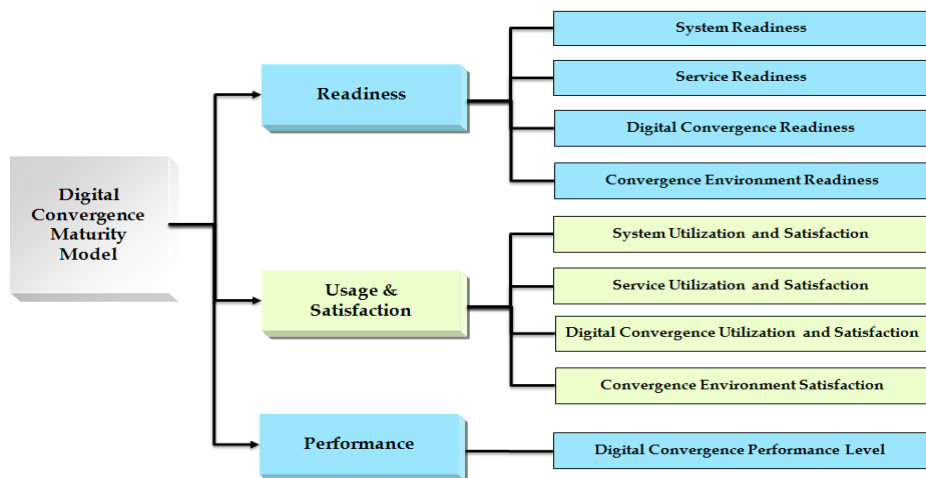


Fig. 1. DCMM Framework

The DCMM, developed in a manner to satisfy these three criteria, proved effective in remedying flaws of existing indices, since scale-type, maturity and impact indices were appropriately combined in this model.

Meanwhile, to determine the level of use of digital convergence services, we designed an index which can allow us to determine the level of use for each field by dividing network and content into five stages each and measuring the current level and required level for each field.

The DCMM so developed, applying the above considerations, is described in the diagram in (Fig.1). The framework was designed drawing on reference materials from the WEFT and IMD, and its main factors are readiness, usage and satisfaction and performance².

Readiness, here, means the readiness of digital convergence services in each field evaluated. Readiness is measured in four subcategories; namely, system readiness, indicating the level of system readiness on the service provider's side, service

² The model, although similar to the NRI framework, is distinct from it in that the constructs used are different. It further distinguishes itself from the NRI in that the results are expressed as levels, and not in specific values, as is the case with the latter framework.

Table 3. Definition of Pillars

Sub Index	Pillar	Definition
Readiness	System readiness	- Level of system provision on the part of the service provider, consisting of five phases: introduction, active use, integration, interlinkage and maturity.
	Service readiness	- Level of interfacing between the service provider and user, consisting of five phases: emerging presence, enhanced presence, interactive presence, transactional presence, and networked presence.
	Convergence readiness	- Level of development in the use environment for digital convergence services, consisting of two perspectives: network and content perspectives. - Each perspective consists of five phases according to the level of network and content development.
	Convergence environment readiness	- Readiness of the convergence environment, measured in five areas: legal and regulatory infrastructure, human capacity, investment, leadership and standardization.
Usage and satisfaction	Usage of and Satisfaction with systems	- Extent to which systems are used and Level of overall satisfaction with the converged digital media service system.
	Usage of and Satisfaction with services	- Extent to which services are used and Level of overall satisfaction with digital convergence services.
	Usage of and Satisfaction with digital convergence	- Extent to which digital convergence services are used and Level of overall satisfaction with digital convergence.
	Satisfaction with convergence environment	- Level of overall satisfaction with the convergence environment
Performance	Converged digital media service-attributable performance enhancement	- Performance benefits realized from the utilization of digital convergence services, consisting of five phases: cost cut, work improvement, promotion of collaborative work, innovation in medical consultation process and creation of new business models.

readiness, indicating the level of interfacing between the service provider and user, digital convergence readiness, indicating the level of maturity of the use environment for digital convergence services, and convergence environment readiness, indicating the readiness for digital convergence services at a societal level. Usage-and-satisfaction refers to the level of use of, and level of satisfaction with, digital convergence services in the four fields studied. Usage-and-satisfaction is measured in four sub-categories, including system utilization and satisfaction, service utilization and

satisfaction, digital convergence utilization and satisfaction and convergence environment satisfaction. Performance means performance benefits resulting from the utilization of digital convergence services in each of the four fields evaluated. The focus was on the level of contribution by digital convergence services to performance in high-level tasks. Details are given in Table 3 above:

Here, for the stages in system readiness, the stages of development for general information systems were used, while in the case of service readiness, the stages were designed drawing on the UN's web access model.

3 Survey Design and Analysis Results

To determine main factors contributing to the promotion of digital convergence services, in this study, we measured the relative importance of readiness, usage and satisfaction and performance in four fields in which digital convergence services are currently used (education, healthcare, transportation, public administration).

To do so, we conducted a survey of 80 experts each specialized in one of the four fields. For the selection of experts, we employed the technique of purposive sampling, using a list, as well as the snowball sampling method. Table 4 provides the summary description of the survey performed for this study:

Table 4. Survey Design Summary

Target sample	Experts in the fields of education, healthcare, transportation and public administration
Method of survey	Delphi survey (using email and fax)
Sample size	80 total experts (education: 19, healthcare: 20; transportation:18, public administration: 23)
Sampling method	Purposive sampling using a list (purposive sampling), Snowball Sampling ³
Survey period	Oct. 15, 2010 – Feb. 11, 2011

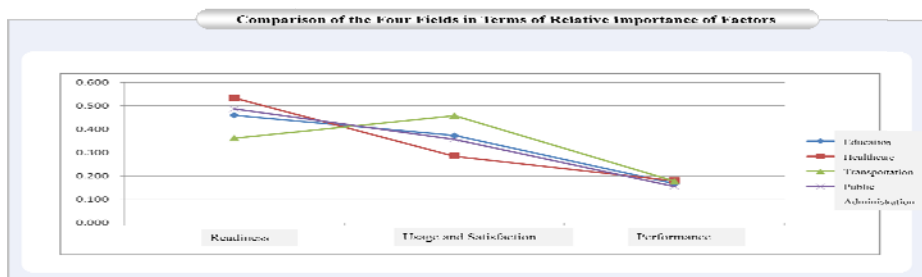


Fig. 2. Comparison of Sub-indices by Field for Relative Importance

³ Snowball Sampling is a technique used to identify potential subjects in studies where subjects are hard to locate. The researcher asks a small number of initial subjects to help identify people with a similar trait of interest.

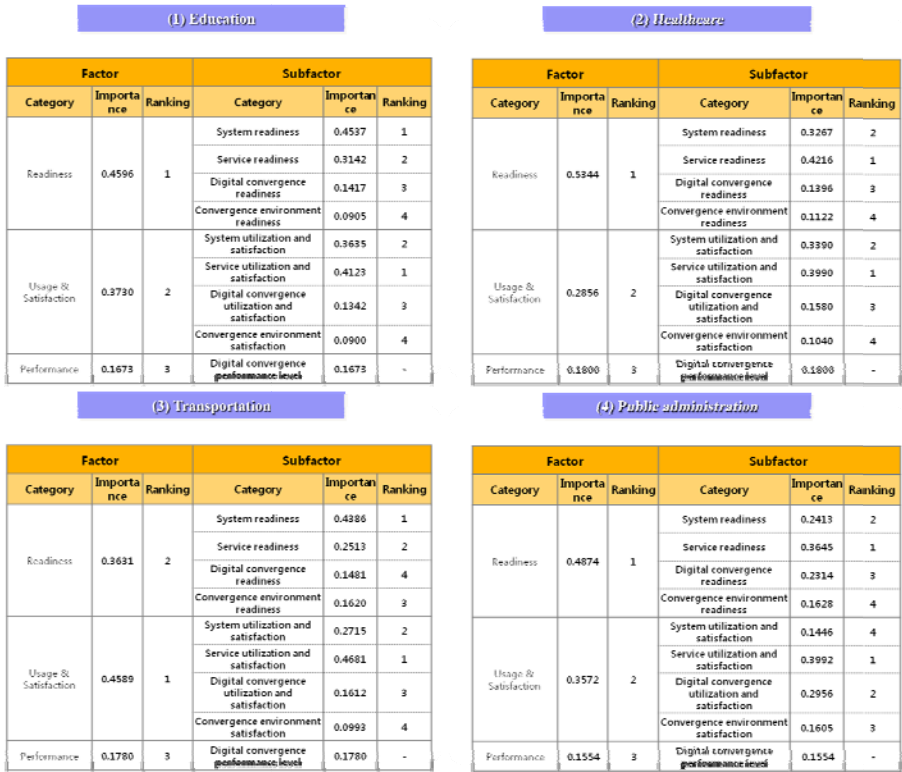


Fig. 3. Relative Importance of Pillars in the Four Fields

Based on the results of the above-mentioned survey, in terms of the relative importance of the sub-indices for the four fields, readiness was the most important in education and healthcare, followed by usage and satisfaction, and performance, in this order. In the transportation field, usage and satisfaction proved the most important, followed by readiness and performance, in this order. In public administration, readiness was the most important, and usage/satisfaction and performance were second and third most important. It is noteworthy that, in this study, readiness and usage proved more important than performance, considered to retain central importance from the perspective of information systems. These results may owe to the fact that fields evaluated in the present study are public-interest fields, rather than for-profit fields. Also experts evaluate that these services are still remained relatively lower stages in maturity. And it may show that supplier's perspective is considered relatively stronger than customer's perspective.

An additional analysis was conducted, meanwhile, to determine the relative importance of the pillars composing the sub-indices as well. The results of this analysis are as shown in (Fig. 3). Concerning readiness, system readiness appeared the most important in the fields of education and transportation, and service readiness in the fields of healthcare and public administration. These results may show that experts are

considered more important internal customers in service organization in the fields of education and transportation.

4 Conclusion

This study, aimed at determining key factors for the promotion of digital convergence services, measured the relative importance of readiness, usage and satisfaction and performance in four fields in which these services are currently in use; namely, education, healthcare, transportation, public administration. The data were obtained through a survey of 80 experts specialized in the four fields, selected using the purposive sampling technique based on a list and also the snowball sampling technique. The AHP was used for analysis. The analysis showed that in the fields of education and healthcare, readiness was the most important, followed by usage/satisfaction, and performance, in this order. In the field of transportation, usage and satisfaction mattered the most, followed by readiness and performance, in this order. In public administration, readiness was of highest importance, and usage/satisfaction and performance were second and third most important. These orders of importance can be useful references in the establishment of future digital convergence-related policies, as they are indicators of what the policy priorities should be.

References

1. Economist Intelligent Unit: Foresight 2020: Economic, Industry and Corporate trends (2007)
2. Kang, S., Kim, J.: The Advent of the u-Health Age. In: SERI CEO Information (2007)
3. Korea Communications Commission: A Strategy for the Development of Converged Media-type Content (2009)
4. Korea Communications Commission: Korea Communications Commission Annual Report (2009)
5. Korea Institute for Industrial Economics and Trade: Industry Vision 2030 (2006)
6. Korea Institute of Science and Technology Evaluation and Planning: Korea's Science and Technology Vision 2040 (2010)
7. Ministry of Education and Human Resources: The History of Education Innovation through e-Learning (2008)
8. Ministry of Health and Welfare: u-Healthcare Survey (2007)
9. Oh, J.: The Current Progress in Healthcare Informatization and Future Tasks, NCA CIO REPORT, National Computerization Agency (2006)
10. Samsung Economic Research Institute: The Economic Effect of u-Health and Development Strategy. SERI Issue Paper (2007)
11. Song, T.: u-Health Trends in Japan, Korea Institute for Health and Social Affairs (2009)
12. Wetering, R., Batenburg, R.: A PACS Maturity Model: A Systematic Meta-Analytic Review on Maturation and Evolvability of PACS in the Hospital Enterprise. *International Journal of Medical Informatics* 78, 127–140 (2009)

Examining Consumer Preferences for Mobile Devices to Read Digital Content for the Diffusion of Ubiquitous Learning in Higher Education

Hyeongjik Lee¹, Won Bin Lee², and Soo Cheon Kweon¹

¹ Technology Strategy Research Division, Electronics and Telecommunications Research Institute, 218 Gajeongno, Yuseong-gu, Daejeon, 305-700, Korea

² Department of Management Science, Republic of Korea Naval Academy, 88-1 Angok-dong, Jinhae, Kyungnam 645-797, Korea
chandler@etri.re.kr, skybini@hotmail.com, sckweon@etri.re.kr

Abstract. We empirically investigate consumer preferences for mobile devices to read digital textbooks and other educational content for the diffusion of ubiquitous learning in higher education by a conjoint method. The results show that current forms of tablet PCs (also called smart pads) might not be suitable for an education device in terms of screen size and type of platform. Thus, policy makers need to consider developing educational mobile devices for the successful diffusion of ubiquitous learning in the long term.

Keywords: educational mobile device, ubiquitous learning, higher education, consumer preference, conjoint analysis.

1 Introduction

The rapid development of Information and Communication Technologies (ICT) has had an impact on the way we teach and learn [1, 2] and various types of devices and equipment have been introduced in education. For example, one element of a smart classroom [3], a smart board, which is a computer-based touch screen and a digital projector that conveys the image on a computer screen to the touch screen, has been developed as a powerful teaching tool for the classroom [4, 5]. From the end-users' perspective, mobile devices have a tremendous potential for use in education [1, 6]. They can directly provide ubiquitous learning environments for students because they can be used to deliver digital textbooks and other education content to students anywhere and anytime [2, 7-10]. Therefore, mobile devices can effectively contribute to the early growth of ubiquitous learning in education although there are several obstacles to the use of mobile devices, including small screens and limited computational power [7, 11, 12].

In spite of the importance of the roles mobile devices can play in ubiquitous learning in education, there has not yet been empirical research conducted on identifying essential features of mobile devices as educational tools. Consequently, this paper empirically investigates the optimal characteristics of mobile devices in

higher education using an analysis of consumer preferences. Consumer preference may be useful in determining the essential features that mobile devices should have because a large part of the development of IT products has been dominated by the pull of demand rather than the push of technology [13]. We use a conjoint analysis to empirically examine consumer preferences for mobile devices as an essential element for ubiquitous learning in higher education.

The remainder of the paper is organized as follows: Section 2 briefly analyzes the essential characteristics of mobile devices for ubiquitous learning in higher education. Section 3 then empirically examines consumer preference for mobile devices in higher education by a conjoint method. Some meaningful implications for policy makers are suggested in section 4. The final section offers our conclusion.

2 The Essential Features of Mobile Devices as Educational Tools

Ubiquitous learning is usually defined as an education system that uses the technologies of ubiquitous computing, wireless communication, mobile devices and context-aware technologies in an education context [14, 15]. As one of the essential elements in ubiquitous learning, mobile devices play important roles in delivering learning materials and other education content anywhere and anytime [2, 7-10]. As there are some features for ubiquitous learning including permanency, accessibility, immediacy, interactivity, and situating of instructional activities [14-16], mobile devices for ubiquitous learning should have some essential values as explained below.

First, educational mobile devices should allow students to freely choose the time and location of study [6, 10]; in other words, they should ensure ‘portability’ as much as possible. This requires that the devices have not only wireless data access technology, including advanced cellular technologies, but also appropriate size and weight. Second, mobile devices should have ‘content display’ adequate enough to easily read digital textbooks and other content. Different from an e-Book reader, which usually displays text-based content, mobile devices for ubiquitous learning should have advanced functions to deliver various types of multimedia-based educational content to students. To do this, educational mobile devices should have screen displays large enough to overcome the clear limitation that current smart phones have. Third, educational mobile devices should also have sufficient functions to make office documents and other educational content using some kind of office applications; in other words, they should assure proper ‘content productivity.’ Most students usually want to read and make document files using office applications such as Word, PowerPoint and Excel before, during, and after class.

Based on these requirements for ubiquitous learning, we briefly analyzed three different forms of mobile devices for ubiquitous learning in higher education: laptops, smart phones, and tablet PCs. First, a notebook computer is a general device used by many college students for educational purposes. Its screen size is usually larger than 12 inches and most of them use a Windows platform and are supported by a fixed internet and Wi-Fi network. While a notebook computer may have advantages in content display and content productivity it has limits in portability.

Second, a smart phone can be expected to be a promising education tool because it combines telephone capability with a PDA, camera, video, mass storage, MP3 player, Internet access and networking features in one compact system [10, 11]. A large number of students have been already using smart phones [6] thus the smart phones have become an icon of a mobile convergent device [17]. For example, by both the increasing popularity of e-books and the willingness of people to read materials on their phones, smart phones may be popular for delivery of instructional materials [17]. However, because their small screen size has difficulty in web browsing and reading text, the use of smart phones in higher education has not yet expanded. In other words, in spite of its high portability, a smart phone has a disadvantage in meeting consumers' needs in terms of content display and content productivity.

Third, a tablet PC, also called a "smart pad," can be also be expected to be a promising educational tool in higher education because it not only has the advantages of smart phones but it also overcomes their limitation. It usually has a larger screen than smart phones even though several different screen sizes of tablet PCs have been introduced. Thus, it has the main function of e-book readers, that is, to display instructional materials, along with the other various functions that PCs and smart phones currently offer. In other words, different from laptops and smart phones, tablet PCs can provide all the value offerings at an appropriate level by filling the gaps between notebook computers and smart phones.

3 Empirical Analyses

3.1 Method: Conjoint Analysis

We empirically examine consumer preferences for mobile devices for ubiquitous learning in higher education using a conjoint approach. The conjoint analysis method is generally used to understand the importance of different product components or product features [12, 18-21]. As conjoint analysis has become an increasingly popular approach to estimating the benefits received from the attributes of a product [20], it has widely used quantitative tools not only in marketing research [21] but also in ICT technology [12, 13, 17, 19] and ICT services [20-24].

To examine the consumer preference for mobile devices as one of essential elements for ubiquitous learning in higher education, we designed and conducted a survey asking respondents to rank a set of alternatives. In the survey, four valuable attributes were selected, and this paper sets some levels for each for a conjoint analysis to determine the relative importance of the attributes and preferences for the levels among the users. The first attribute is the screen size of mobile devices. Screen size is the most critical factor determining the form of the mobile device [13] and a small screen display is usually recognized as a clear limitation of a mobile device for mobile learning in higher education [7, 12]. Because mobile devices for mobile learning should meet the requirements of readability and legibility [7], in terms of content display, we consequently considered the following three levels: 7 inches (one favorite size of tablet PCs), 10 inches (another favorite size of tablet PCs), and 12 inches (the size of small laptops).

The second attribute is the type of platform, which usually includes the operating system (OS) of a mobile device. Different from a laptop whose platform is based on Windows, smart phones include both a web browser, which provides access to the wealth of material on the World Wide Web and inexpensive applications (usually called apps) based on mobile platforms such as Google's Android [6]. In other words, compared with the traditional PC based platforms including Windows, mobile platforms provide different user interfaces and user experiences for end-users. We then considered the following two levels: a PC-based platform (usually a Windows platform) and a mobile-based platform (such as Apple's iOS and Google's Android).

The third attribute is the level of office productivity which mobile devices provide. Since most students read and make document files using office applications such as Word, PowerPoint and Excel, in terms of content productivity, the level of functions related to office applications can importantly affect the preference for mobile devices in higher education. From the perspective of hardware, input equipment may be directly related to office productivity. A keyboard with a trackball mouse is the most preferred input equipment but it has limits in portability, while a touch-screen, which is the main input equipment of smart phones and tablet PCs, has disadvantages in content productivity [13]. From the view of mobile devices' software, the function of office applications in mobile platforms has some limits compared to that of PC platforms. We consequently suggested the following three levels: high level (easy to make and read document files), medium level (easy to read but not so easy to make document files), and low level (only read document files).

The last attribute is the wireless access technology that mobile devices offer. To use most mobile content through mobile devices, a data connection with mobile networks using data access technology is necessary. Currently, most mobile devices support Wi-Fi technology and some of them can access a cellular network. We then considered the following two levels: support for Wi-Fi technology only and support for a combination of both Wi-Fi and a cellular network.

Considering the four attributes shown in Table 1, this study generates 36 (=32×22) profiles based on a full profile method. Because respondents clearly have difficulty in completing all profiles, we use fractional factorial design (FFD), which simplifies the number of profiles to be tested and maintains the effectiveness of sorting and evaluating the relative importance of a product's multi-dimensional attributes [13, 19, 20]. Based on FFD, the number of initial profiles is reduced to 9 profiles. After one unrealistic and inappropriate combination among them is excluded, the final 8 profiles were used to conduct a survey.

Table 1. Attributes and their levels

Attribute	Level
Screen Size	7 inches / 10 inches / 12 inches
Platform	Mobile-based / PC-based (Windows)
Office Productivity	High / Medium / Low
Wireless Technology	Wi-Fi Only/ Wi-Fi + Cellular

3.2 Results

To examine consumer preference for mobile devices for ubiquitous learning in higher education, we carried out a survey asking respondents to rank a set of alternatives. Each respondent was asked to rank the 8 profiles describing the form of mobile devices as an education tool according to their own usage intention on a scale from 1 (most preferred) to 8 (least preferred). Respondents were selected in an “A” class national university in South Korea because all the students in the university had received a laptop as their educational devices without any cost. Therefore, respondents could evaluate the profiles regardless of the price of mobile devices. In addition, most college students usually have more knowledge about mobile devices. Thus, we could obtain more meaningful results from the respondents who were the main users of mobile devices as an education tool without consideration of the price of the devices.

A total of 202 respondents were interviewed from September 5 to September 16 in 2011 via a Web page. Among these respondents, we excluded 54 respondents who had missing values. Thus, the data of the final 148 respondents were used for analysis. Among them, 132 (89.2%) were males and 16 (10.8%) were female. 139 respondents used smart phones (93.9%) and 15 respondents additionally used Tablet PCs (10.1%).

Table 2 shows the results of the conjoint analysis. The results show that respondents importantly considered all the attributes since the relative importance of each attribute was higher than 15%. The most important attribute was screen size (relative importance 31.1%). Surprisingly, respondents preferred mobile devices with 12 inch screens, which current tablet PCs do not have. This implies that respondents want to receive digital education content, including digital textbooks and lecture materials, on a large screen display even though they have difficulty in portability.

The next important attribute was the type of platform (relative importance 26.3%). The result shows that respondents preferred the Windows platform to mobile platforms. This means that most respondents are accustomed to PC-based platforms

Table 2. Conjoint results

Attribute	Level	Utility Estimate	Standard Error	Relative Importance
Screen Size	7 inches	-0.467	0.087	31.17
	10 inches	0.055	0.087	
	12 inches	0.412	0.103	
Platform	Mobile-Based	-0.402	0.086	26.23
	Windows	0.402	0.086	
Office Productivity	Low	-0.248	0.103	25.94
	Medium	0.492	0.112	
	High	-0.244	0.087	
Wireless Technology	Wi-Fi Only	-0.282	0.077	16.66
	Wi-Fi + Cellular	0.282	0.077	
Constant	-	4.885	0.085	-

Pearon’s R=0.992 (0.000), Kendall’s tau=1.000 (0.000).

(Windows) and thus want to use mobile devices with a platform consistent with the platform of their PCs at home or at their university, even though they have been enjoying their time using smart phones with platforms that are mobile-based.

The results also show that respondents prefer a medium level of office productivity (relative importance 26.0%). This implies that they don't need comprehensive functions when they make office documents and other content. In terms of wireless technology (relative importance 16.7%), most respondents want to receive mobile content using Wi-Fi and cellular technology. In other words, respondents want widespread service coverage; thus cellular technology is principally required for the mobile devices for ubiquitous learning in higher education.

4 Implications of the Empirical Findings for Policy Makers

We examined which features mobile devices should have for ubiquitous learning in higher education using a conjoint analysis and the results show that respondents preferred large screen, Windows platform based mobile devices as their educational tools. This paper also has some interesting and meaningful implications for policy makers to understand the requirements of mobile devices for the successful growth of ubiquitous learning in higher education.

First, in terms of screen size, the result indicates that current tablet PCs may not be optimal mobile devices for ubiquitous learning because potential customers may not be satisfied with 7 inches or 10 inches mobile devices. Therefore, policy makers need to consider developing and introducing specialized mobile devices for ubiquitous learning in higher education in the long term. Considering the demand for ubiquitous learning in higher education, specialized mobile devices for educational purposes may contribute the growth of the ICT industry in South Korea.

Second, in terms of the type of platform, the result shows that potential customers may require a consistent platform regardless of the device used. This implies that education content should be equally delivered through one single platform-based N-screen devices and Microsoft may have advantages in planning N-screen strategy based on Windows platform because the Windows platform can be suitable for education devices. Not surprisingly, Microsoft recently announced the latest integrated platform called 'Windows 8' which supports both PCs and mobile devices, including tablet PCs. Therefore, in the short term, policy makers might consider the introduction of Windows-based tablet PCs for ubiquitous learning although it can have some limits in screen size.

Lastly, the results also imply that, compared with laptops and smart phones, tablet PCs still have the potential as the second-best mobile device for ubiquitous learning in higher education, in spite of its limitations. Although potential customers usually want a larger screen size than tablet PCs currently have, they also need to receive education content anytime and anywhere using cellular networks. Thus portability may still be necessary for mobile devices in higher education, and tablet PCs can be the second best device for ubiquitous learning. Besides, emerging mobile devices that blur the line between smart phones and tablet PCs have been continuously introduced.

Therefore, policy makers need to keep reviewing several alternatives and finding the optimal one among them rather than waiting for the best mobile device for ubiquitous learning.

5 Conclusion

As mobile devices can be used to deliver digital textbooks and other education content to students anywhere and anytime, they can effectively contribute to the early growth of ubiquitous learning in education although they have some limitations, such as small screen displays and limited computational power. We investigated the essential requirements of mobile devices for ubiquitous learning by using a consumer preference analysis. The results showed that, in spite of the expectation, tablet PCs might not be optimal for ubiquitous learning in terms of screen size and type of platform. We also found that potential customers do not want to use comprehensive functions related to office documentation through their mobile devices while they want to receive ubiquitous access by using advanced cellular networks. Therefore, We suggest that policy makers need to consider developing and introducing a specialized mobile device for ubiquitous learning in the long term, but for the early growth of ubiquitous learning, they also need to consider introducing Window-based tablet PCs as the second-best alternative in the short term.

Although we presented some meaningful findings, it has also some limitations. First, it examined consumer preferences for mobile devices in higher education by considering four features. To fully understand the requirements of mobile devices for ubiquitous learning, further research should examine other characteristics that mobile devices should have. Second, by limiting the focus of this paper to South Korea, the generality of the empirical results should be treated with caution. Therefore, a useful area of future research would be to extend the empirical analysis to other samples or to an international context.

References

1. Martin, S., Diaz, G., Sancristobal, E., Gil, R., Castro, M., Peire, J.: New technology trends in education: Seven years of forecasts and convergence. *Computers & Education* 57, 1893–1906 (2011)
2. Kalhor, Q., Chowdhry, L.L., Abbasi, T., Abbasi, S.: M-learning-an innovative advancement of ICT in education. In: 2010 4th International Conference on Distance Learning and Education (ICDLE), pp. 148–151. IEEE (2010)
3. Sevindik, T.: Future's learning environments in health education: The effects of smart classrooms on the academic achievements of the students at health college. *Telematics and Informatics* 27, 314–322 (2010)
4. Gursul, F., Tozmaz, G.B.: Which one is smarter? Teacher or Board. *Procedia - Social and Behavioral Sciences* 2, 5731–5737 (2010)
5. Al-Qirim, N.: Determinants of interactive white board success in teaching in higher education institutions. *Computers & Education* 56, 827–838 (2011)

6. Williams, A.J., Pence, H.E.: Smart Phones, a Powerful Tool in the Chemistry Classroom. *Journal of Chemical Education* 88, 683–686 (2011)
7. Jun, Y., Zhi-yi, H.: Learning anywhere anytime with mobile devices. In: *International Conference on Artificial Intelligence and Education (ICAIE)*, pp. 768–771. IEEE (2010)
8. Liu, Y., Li, H., Carlsson, C.: Factors driving the adoption of m-learning: An empirical study. *Computers & Education* 55, 1211–1219 (2010)
9. Kumar, L.S., Jamatia, B., Aggarwal, A.: Mobile phones as effective learner support devices: A case study. In: *2010 4th International Conference on Distance Learning and Education (ICDLE)*, pp. 211–213. IEEE (2010)
10. Junfeng, W.: Mobile learning—A new type of distance education. In: *2010 International Conference on Educational and Information Technology (ICEIT)*, pp. 149–152. IEEE (2010)
11. Economides, A., Grousopoulou, A.: Students' thoughts about the importance and costs of their mobile devices' features and services. *Telematics and Informatics* 26, 57–84 (2009)
12. Jeon, H.J., Kim, M.S., Sohn, S.Y.: Conjoint and WTP analyses of future mobile phones for digital convergence. *Technological Forecasting and Social Change* 77, 457–465 (2010)
13. Kim, Y., Lee, J.-D., Koh, D.: Effects of consumer preferences on the convergence of mobile telecommunications devices. *Applied Economics* 37, 817–826 (2005)
14. Tsai, P.-S., Tsai, C.-C., Hwang, G.-H.: College students' conceptions of context-aware ubiquitous learning: A phenomenographic analysis. *The Internet and Higher Education* 14, 137–141 (2011)
15. Wang, S.-L., Wu, C.-Y.: Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications* 38, 10831–10838 (2011)
16. Shih, J.-L., Chu, H.-C., Hwang, G.-J.: An investigation of attitudes of students and teachers about participating in a context-aware ubiquitous learning activity. *British Journal of Educational Technology* 42, 373–394 (2011)
17. Yoo, J., Yoon, Y., Choi, M.: Importance of positive reputation for Smartphone adoption. In: *2010 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 314–318. IEEE (2010)
18. Nam, C., Kim, S., Lee, H.: The role of WiBro: Filling the gaps in mobile broadband technologies. *Technological Forecasting and Social Change* 75, 438–448 (2008)
19. Min, S.H., Kim, H.Y., Kwon, Y.J., Sohn, S.Y.: Conjoint analysis for improving the e-book reader in the Korean market. *Expert Systems with Applications* 38, 12923–12929 (2011)
20. Song, J., Jang, T., Sohn, S.Y.: Conjoint analysis for IPTV service. *Expert Systems with Applications* 36, 7860–7864 (2009)
21. Kargin, B., Basoglu, N., Daim, T.: Exploring mobile service adoption: A conjoint model. In: *Portland International Conference on Management of Engineering & Technology, PICMET 2008*, pp. 2623–2633. IEEE (2008)
22. Nakamura, A.: Estimating switching costs after introducing Fixed-Mobile Convergence in Japan. *Information Economics and Policy* 23, 59–71 (2011)
23. Jeong, G., Koh, D., Lee, J.: Analysis of the Competitiveness of Broadband over Power Line Communication in Korea. *ETRI Journal* 30, 469–479 (2008)
24. Ahn, J., Lee, J., Lee, J.-D., Kim, T.-Y.: An Analysis of Consumer Preferences among Wireless LAN and Mobile Internet Services. *ETRI Journal* 28, 205–215 (2006)

A Service Quality Model for the Public Information Service

Hyun-Woong Jo and Sang-Wook Kim

Chungbuk National University,
#52 Naesudong-ro, Heungduk-gu, Cheongju, Chungbuk 361-763, South Korea
hektor.jo@gmail.com, sierra@cbnu.ac.kr

Abstract. This paper describes the development of a 22-items instrument for assessing customer perceptions of service quality about public information service. Responses from 702 users were obtained in four public information service. Through an evidence of the scale's reliability, factor structure, and validity on the basis of analyzing data from four samples, this study found out there is no problem in measuring the quality of service with the proposed factors and indices.

Keywords: Service Quality, SERVQUAL, Public Information Service.

1 Introduction

Public Information Project in Korea has been constantly carried out since 2001, but considering how e-government has been progressed so far, some criticize that the improvement of governmental operation systems and functions has not surpassed the critical mass to provide efficient administrative services. Since the process was focused on individual organization-centered computerization rather than people-centered service from a technological viewpoint as well as each agency's work process automation rather than pan-governmental work innovation, it is mainly pointed out that the project has failed to extensively lead the ways of leading governmental work in an innovative level.

What is more concerned about is that the investment into information systems is leading to IT Productivity Paradox that causes a declination of organizational achievements instead. As shown in Table 1 below, information system (technological) performances of the e-governmental system is evaluated as considerably superior, but such a technological potential hasn't led to organizational performances (governmental transparency, work efficiency and national competitiveness).

These evaluations don't have direct statistics about the e-government, but it is clear that the achievements of the information system realized under 'the embodiment of e-government' have rarely contributed to governmental organizations' making ultimate achievements such as governmental efficiency and transparency and national competitiveness.

Recently, as people become more aware of the importance of service quality and quality management all over the society, and people have more interest in service

quality and quality management in the field of administrative service, such as e-government, it is needed to convert the basic viewpoint for public information projects from a supplier-centered one with technology push to a demander-centered one with social demand pull.

Table 1. IS Performance vs. Organization Performance

		'02	'03	'04	'05	'08	'10
Information System Performance	e-Government Readiness Index	15	13	5	5	6	1
	National Competitiveness	29	32	31	27	31	23
Organization Performance	Government Efficiency	25	37	36	31	37	26
	Corruption Perception Index	40	50	47	40	40	39

Then, what is a practical alternative to establishing a service delivery system and controlling information service from not a supplier's point of view but a user's point of view? The answer to this question can be found in the evaluation system. To lead the information system project managed in the aspect of service, it is the most practical alternative to measure and evaluate the quality of information service from a user's point of view in terms of the evaluation system. Thus, it is more important than anything else to develop a measurement model and an evaluation index used to objectively evaluate the quality of public information service and constantly improve them.

Based on the awareness of such a problem, therefore, this study attempted (1) to develop a reliable measurement index for the information service quality model in order to secure high quality e-governmental service, (2) to carry out a survey for users about 4 public information systems selected, and (3) to verify the validation of the developed model and its applicability.

2 Theoretical Foundations

Measuring the quality of public information systems is in the same context as achieving information system success, and as a means to attain this goal, this study attempted to suggest a quality measurement model in the aspect of service. Accordingly, the ground for establishing a public information service quality model suggested by this study can be found in 'IS Success Model' and 'Service Quality Model'.

2.1 Information System Success

DeLone and Mclean combined about 180 different researches previously carried out in order to simplify and organize variables that may influence the successful operation of information system. Main variables are categorized into six major domains, which are, system quality, information quality, use, user satisfaction, individual impact, and organizational impact. Once the variables are categorized into different domains, their correlations with other variables are technically modeled, which is the DM IS Success

Model that they suggested [1]. In addition to the empirical researches carried out using this very model, elements such as service quality is added in later stage as well as the user's intention or user satisfaction that may affect the user himself/herself and even the society as a whole to ultimately analyze the net benefit of the system to extract the correlation model as it is shown in figure 1 [2].

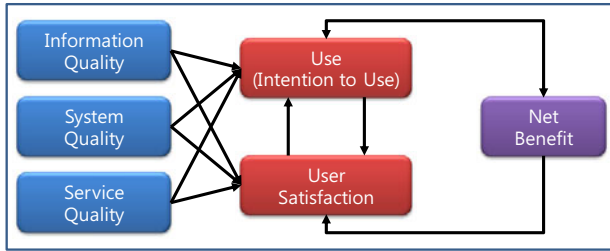


Fig. 1. IS Success Model

2.2 Service Quality

The construct of quality as conceptualized in the services literature and as measured by SERVQUAL involves perceived quality. Perceived quality is the consumer's judgment about an entity's overall excellence or superiority [3].

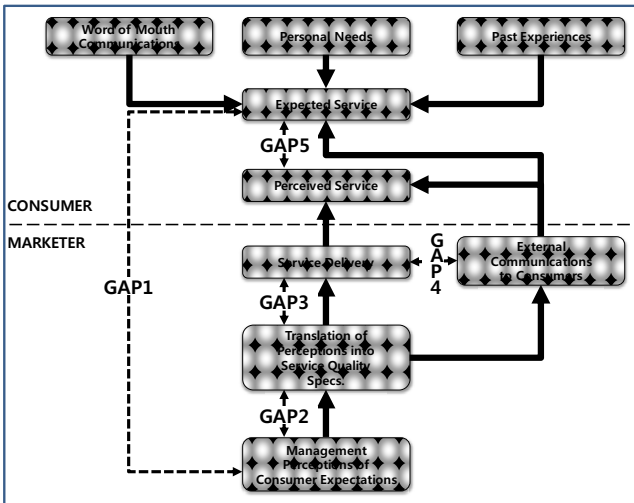


Fig. 2. PZB's Service Quality Model (GAP Model)

It differs from objective quality [4]; it is a form of attitude, related but not equivalent to satisfaction, and results from a comparison of expectations with perceptions of performance [5].

SERVQUAL was built on the theory that service quality is the gap between what customers expect and what performance they actually perceive [5]. Service quality is calculated with SERVQUAL by subtracting expectation scores from perceived performance scores. Respondents are first asked to rate their expectations of service from an excellent firm or organization, and then to rate the performance they perceive from a specific organization.

3 Methodology

IS Success Model proposed 3 factors that affect IS success: Information Quality, System Quality, Service Quality. These are dimensions that make a service quality. Table 2. shows the items about each dimensions [1][6]. Service Quality factor replaces Work System Quality that is about service provider’s quality.

Table 2. Dimensions, Items, Codes

Dimen.	Item	Code	Dimen.	Item	Code	Dimen.	Item	Code
Information Quality	Completeness	IQ1	System Quality	Up-to-Date	SQ1	Work System Quality	Rapidity	WQ1
	Currency	IQ2		Interface	SQ2		Rationality	WQ2
	Accuracy	IQ3		Stability	SQ3		Assurance	WQ3
	Relevance	IQ4		Adaptability	SQ4		Trustworthiness	WQ4
	Flexibility	IQ5		Functionality	SQ5		Empathy	WQ5
	Readability	IQ6		Compatibility	SQ6		Sensitivity	WQ6
				Availability	SQ7		Appearance	WQ7
				Responsiveness	SQ8			
				Security	SQ9			

4 Reliability and Factor Structure

4.1 Data Collection

Data of the 22-item instrument were gathered from a sample of 702 users. Screened and qualified respondents self administered a two-part questionnaire consisting of a 22-statement expectations part followed by a 22-statement perceptions part. For the first part, respondents were instructed to indicate the level of service that should be offered by information service. For the second part, respondents were then instructed to express their perceptions about information service.

4.2 Reliability Analysis

The raw data used in computing coefficient alpha were in the form of difference scores. Specifically, for each item a difference score Q was defined as $Q = P - E$, were P and E are the ratings on the corresponding perception and expectation statements, respectively. The value of coefficient alpha ranged 0.953, 0.927 across the 2 dimensions.

Table 3. Survey Results

		A	B	C	D
	Total	246	240	79	137
Gender	Male	165	198	56	93
	Female	80	42	23	42
Job	Public Official	241	110	11	82
	Specialized Job	-	130	7	-
	Citizen	3	-	61	55
Use Period	~1Month	17	3	6	15
	1~6Months	14	17	20	25
	0.5~1Year	39	46	14	21
	1~2Years	59	106	19	36
	2Years~	117	68	20	40

Examining the dimensionality of the 22-item scale was the next task in this stage of scale purification and was accomplished by factor analyzing the difference scores on the 22-items. The principal axis factoring procedure [7] was used and the analysis was constrained a priori to 3 factors. When the 3-factor solution was rotated orthogonally, no clear factor pattern emerged. Many of the items had high loadings on several factors, thereby implying that the factors may not be independent of one another. Therefore the 3-factor solution was subjected to oblique rotation to allow for intercorrelations among the dimensions and to facilitate easy interpretation.

Factor analysis was resulted in a pool of 22-items representing 2 distinct dimensions.

The high alpha values indicated good internal consistency among items within each dimension. Moreover, the combined reliability for the 22-item scale, computed by using the formula for the reliability of linear combinations [8], was quite high (0.965).

Table 4. Reliability Analysis

Dimension	Label	Cronbach's Alphas	Number of Items	Item Code
Information Quality	F1	0.953	15	IQ1~6
System Quality				SQ1~9
Work System Quality	F2	0.927	7	WQ1~7
Total Scale Reliability (Reliability of Linear Combination)		0.965	22	-

4.3 Internal Consistency Analysis

Table 6 shows the component and total reliabilities for each of the four samples. The reliabilities are consistently high across all four samples. The total scale reliability is close to 0.9 in each of the four instances.

Results of the factor analyses of data from the four samples are summarized in Table 7. The overall patterns of factor loadings are remarkably similar across the four independent sets of results.

Table 5. Factor Analysis

Dimension	Measure	Code	F1	F2
Information Quality	Completeness	IQ1	0.842	-0.079
	Currency	IQ2	0.727	0.033
	Accuracy	IQ3	0.638	0.106
	Relevance	IQ4	0.791	-0.031
	Flexibility	IQ5	0.846	-0.093
	Readability	IQ6	0.772	0.017
System Quality	Up-to-Date	SQ1	0.761	0.061
	Interface	SQ2	0.791	-0.002
	Stability	SQ3	0.785	0.054
	Adaptability	SQ4	0.765	0.025
	Functionality	SQ5	0.806	0.017
	Compatibility	SQ6	0.804	-0.039
	Availability	SQ7	0.752	-0.017
	Responsiveness	SQ8	0.694	0.139
	Security	SQ9	0.526	0.265
Work System Quality	Rapidity	WQ1	0.303	0.575
	Rationality	WQ2	0.266	0.611
	Assurance	WQ3	0.135	0.734
	Trustworthiness	WQ4	0.066	0.781
	Empathy	WQ5	-0.049	0.901
	Sensitivity	WQ6	-0.065	0.932
	Appearance	WQ7	-0.059	0.840

Table 6. Internal Consistency Analysis

Dimension	Label	Cronbach's Alphas				Number of Items
		A	B	C	D	
Information Quality	F1	0.886	0.956	0.965	0.935	IQ1~6
System Quality						SQ1~9
Work System Quality	F2	0.939	0.930	0.936	0.882	WQ1~7
Total Scale Reliability (Reliability of Linear Combination)		0.930	0.969	0.975	0.950	22

Table 7. Factor Loading Matrices

Item	Factor Loadings								
	A		B		C		D		
	F1	F2	F1	F2	F1	F2	F1	F2	
IQ	1	0.724	-0.027	0.786	0.065	0.840	0.033	0.595	0.286
	2	0.723	-0.063	0.767	0.040	0.475	0.390	0.812	0.322
	3	0.623	-0.111	0.767	0.030	0.437	0.405	0.816	0.169
	4	0.651	-0.192	0.995	-0.228	0.668	0.256	0.754	0.269
	5	0.748	-0.032	0.815	0.028	0.830	-0.051	0.562	0.256
	6	0.713	-0.080	0.637	0.170	0.479	0.448	0.544	0.360
SQ	1	0.770	-0.045	0.647	0.221	0.847	0.052	0.868	0.108
	2	0.788	0.027	0.652	0.197	0.594	0.299	0.756	0.196
	3	0.790	-0.025	0.485	0.422	0.460	0.434	0.483	0.465
	4	0.803	0.003	0.464	0.368	0.639	0.173	0.753	0.164
	5	0.806	-0.022	0.573	0.332	0.729	0.166	0.709	0.237
	6	0.898	0.182	0.452	0.392	0.706	0.122	0.809	0.075
	7	0.842	0.016	0.552	0.268	0.840	-0.003	0.511	0.388
	8	0.870	0.008	0.528	0.342	0.482	0.412	0.539	0.220
	9	0.827	-0.053	0.704	0.057	0.697	0.215	0.658	0.132
WQ	1	0.197	-0.717	0.203	0.655	-0.127	0.994	0.344	0.500
	2	0.190	-0.734	0.113	0.778	-0.081	0.882	0.333	0.686
	3	0.107	-0.779	0.096	0.796	-0.224	1.057	0.400	0.533
	4	0.067	-0.832	-0.100	0.912	0.291	0.611	0.450	0.568
	5	-0.057	-0.923	-0.124	0.922	0.032	0.818	0.347	0.771
	6	-0.120	-0.944	-0.037	0.898	0.147	0.719	0.294	0.768
	7	-0.009	-0.814	0.049	0.687	0.205	0.552	0.238	0.754

5 Assessment of Validity

Research model's high reliabilities and consistent factor structures across several independent samples provide support for its trait validity [9][10]. However while high reliabilities and internal consistencies are necessary conditions for a scale's construct validity, they are not sufficient [11]. The scale must satisfy certain other conceptual and empirical criteria to be considered as having good construct validity.

The scale's validity was also assessed empirically by examining its convergent validity. Respondents rated the information service's overall information system quality (Overall ISQ) and overall work system quality (Overall WSQ) by checking one of 7 scale (1 point: very poor, 7 point : very excellent). The correspondence between the Overall ISQ, Overall WSQ ratings and the research model scores was examined using one-way ANOVA. The treatment variable in the ANOVA's was Overall ISQ, Overall WSQ – with three categories instead of 7 scale: 1, 2 point combined “poor” category, 3~5 point combined “good” category, and 6, 7 point combined “excellent” category.

The results of these analyses for each of the four samples are summarized in Table 8. The strength and persistence of the linkage between the Overall ISQ, Overall WSQ categories and the scores across overall data samples offer strong support for

Table 8. Convergent Validity

1. Overall System (* p<0.05)						
Dimension	Overall ISQ			Overall WSQ		
	Excellent	Good	Poor	Excellent	Good	Poor
IQ, SQ	-0.379*	-0.869*	-1.292*	-0.291*	-0.918*	-1.275*
WQ	-0.199*	-0.755*	-1.408*	-0.075*	-0.815*	-1.450*

2. A System (* p<0.05)						
Dimension	Overall ISQ			Overall WSQ		
	Excellent	Good	Poor	Excellent	Good	Poor
IQ, SQ	-0.651*	-0.781*	-1.539*	-0.584*	-0.798*	-1.722*
WQ	-0.309*	-0.661*	-1.964*	-0.162*	-0.723*	-2.048*

3. B System (* p<0.05)						
Dimension	Overall ISQ			Overall WSQ		
	Excellent	Good	Poor	Excellent	Good	Poor
IQ, SQ	-0.111	-0.751	-0.673	-0.007*	-0.837*	-0.564*
WQ	-0.008	-0.726	-0.540	0.111*	-0.807*	-0.520*

4. C System (* p<0.05)						
Dimension	Overall ISQ			Overall WSQ		
	Excellent	Good	Poor	Excellent	Good	Poor
IQ, SQ	-0.148	-0.991	-4.844	-0.217	-0.942	-3.307
WQ	-0.278	-1.038	-5.333	-0.205	-1.008	-3.886

4. D System (* p<0.05)						
Dimension	Overall ISQ			Overall WSQ		
	Excellent	Good	Poor	Excellent	Good	Poor
IQ, SQ	-0.626	-0.850	-0.974	-0.400	-0.415	-1.051
WQ	-0.467	-0.802	-1.786	-0.295	-0.875	-2.000

convergent validity. And the strength and persistence of the linkage between the Overall ISQ, Overall WSQ categories and the scores across four independent samples offer strong support for convergent validity (except B system).

6 Conclusion

The purpose of this study is to suggest a new way of measuring the quality of information service to order to solve the problem that the information system (technological) performances don't lead to organizational performances. Through a reliability analysis and a feasibility analysis, this study found out there is no problem in measuring the quality of service with the factors and index of IS Success Model.

However, as a product in combination of service quality control in marketing and information system success factors in the filed of MIS on the basis of the existing

theories, the result of this study is merely a generic model. Therefore, it seems needed to conduct an operation of refinement and elaboration for the developed model in the future.

Moreover, since the result of service quality measurement should be practically helpful to information service providers, it is important to suggest guidelines of diagnosis and improvement, so further studies should be conducted on extracting improvements and guidelines for the measurement result through each index.

References

1. DeLone, W.H., McLean, E.R.: Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research* 3(1) (1992)
2. DeLone, W.H., McLean, E.R.: The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems* 19(4) (2003)
3. Zeithaml, V.A.: *Defining and Relating Price, Perceived Quality, and Perceived Value*. Marketing Science Institute, Cambridge (1987)
4. Garvin, D.A.: Quality on the Line. *Harvard Business Review* 61, 65–73 (1983)
5. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing* 64(1) (1988)
6. Seddon, P.B.: A Respecification and Extension of DeLone and McLean Model of IS Success. *Information Systems Research* 8(3) (1997)
7. Harman, H.H.: *Modern Factor Analysis*, 2nd edn. The University of Chicago Press, Chicago (1967)
8. Nunnally, J.C.: *Psychometric Theory*, 2nd edn. McGraw-Hill Book Company, New York (1978)
9. Campbell, D.T.: Recommendations for APA Test Standards Regarding Construct, Trait, or Discriminant Validity. *American Psychologist* 15 (August 1960)
10. Peter, J.P.: Construct Validity: A Review of Basic Issues and Marketing Practices. *Journal of Marketing Research* 18 (May 1981)
11. Churchill Jr., G.A.: A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research* 16 (February 1979)

A Study on the Smartwork Policy for the Productivity in Ageing Advanced Economies

Hee-kyung Kong and Kihoon Sung

Electronics and Telecommunications Research Institute, Daejeon, Gajeongno,
Yuseong-gu, Daejeon, Korea Republic of Korea
{konghk, skh64261}@etri.re.kr

Abstract. The introduction of advanced capabilities in information and communication technology (ICT) has enabled organizations to support work activities and collaborate effectively. Through this usage, ICT has also enabled the introduction of new organizational structures and business process designs that offer previously unavailable flexibility in when and where work is performed. This paper intends to propose a policy for smartwork to be deployed with ICT investment to contribute to increasing labor productivity in the ageing society.

Keywords: Smartwork, Productivity, Ageing, Policy.

1 Introduction

With the ongoing ageing of society, the deterioration of labor productivity has emerged as a major social issue, and smartwork has been recognized as one of the means of resolving the problem. Along with the many discussions about increasing productivity through ICT investment in the past, expectations that smartwork will play a key role in answering the social demand to solve environmental pollution and climate change problems, national demands to save costs and increase efficiency, and personal demands to improve the quality of life and work are increasing. This study intends to propose a policy for smartwork to be deployed with ICT investment to contribute to increasing labor productivity in the ageing society. Chapter 2 discusses the productivity problem expected in the ageing society, Chapter 3 introduces smartwork as the means to solve the problem, Chapter 4 discusses ways of establishing an effective smartwork policy, and Chapter 5 presents the conclusion.

2 Productivity Issues in Ageing Advanced Economies

Extended life expectancy helped by the low birthrate and the advancement of medical technology and systems is causing the rapid ageing of future society. The global fertility rate (children born per woman) was around 5 immediately after World War II, but is now down to 2.5 life expectancy has increased from 50 to 68 during the same period [1]. The ageing of the population is followed by the ageing of the labor

force. As the absolute population of the key labor force between the ages of 25 and 49 decreases, it is likely that that key labor force spectrum will be extended to those aged over 49, a factor which will significantly change labor productivity. First of all, the lower population of the key labor force will cause a decrease in total labor productivity. Furthermore, as there will be less input from the young labor force, the demand for the older segments of the labor force will additionally increase. Since the productivity of the older labor force is relatively lower than that of the key labor force, it will inevitably lower total labor productivity.

There have been many studies of the aged labor force and its productivity. The aged labor force has the characteristics of 1) lower output compared to wages, 2) fewer actual working hours, and 3) a deterioration of physical and technical attainment capability. Although there are ways to mitigate the productivity decrease somewhat by implementing a wage peak policy or flexible work hours in the case of 1) and 2), the case of 3) can be remedied only with direct investment such as training. However, the effectiveness of reinvestment in the aged labor force has not yet been proven, and its validity is questionable since the aged labor force has less capability to gain much from such training. In the end, the lower productivity of the aged labor force will mean there is a limit to how much their performance can be improved.

Until now, the low productivity of the aged labor force has not been identified as a direct reason for productivity deterioration. In Japan, where the percentage of the aged labor force is very high, the number of workers in the labor force aged 55 years or older increased from 22.5% in 1995 to 23.7% in 2002, but the labor productivity per unit of labor input also increased from 85.4 to 94.6 (on the basis of 100 in 2005). This is because the aged labor force is concentrated in primary industries, such as agriculture, whose impact on overall productivity is not high [2 & 3]. In Korea where ageing is occurring almost as rapidly as in Japan, the percentage of aged workers in the overall workforce increased from 14.3% in 1995 to 15.8% in 2002, but the labor productivity per unit of labor input rapidly increased from 62.2 to 88.2 [4]. However, that productivity increase has been slowing down lately. In the case of Japan, the annual productivity increase rate reached 7.8% in 2004, but began to rapidly decrease to reach -1.3% in 2009. Although not as drastic, the figure in Korea also decreased from 9.7% in 2007 to 3.2% in 2009 [2]. Since the level of ageing as well as the socio-economic dynamics and key industries differ from country to country, it may be farfetched to draw any conclusions from the cases of Japan and Korea alone, but it is a definite fact that the lowering of productivity as a result of ageing is becoming a serious problem.

Direct investment is needed to improve productivity. There has been active ICT investment to improve business efficiency since the 1970's. Many studies have shown how ICT investment has affected the increase in productivity. Since visible productivity increases first appeared with the productivity paradox at the end of the 1980's, there have been active discussions of ICT investment since the mid 1990's. Berndt and Morrison (1995) mentioned the inverse tendency between the distribution of personal computers and GDP growth to support the productivity paradox [5]. Hitt and Brynjolfsson (1999) argued that ICT investment acted as a factor with a significant impact on business performance that eventually brought about a productivity increase

[7]. Sircar, Turnbow and Bordoloi (2000); Nordhaus (2001); and Becchetti, Bedoya and Paganetto (2003) have shown that there is a positive correlation between ICT investment and productivity increase [7, 8, 9].

Despite the ongoing arguments, enterprises continue to invest in ICT to increase productivity. However, the type of ICT investment in the aged society is likely to be different from that which has prevailed until now. This is because the productivity decrease due to ageing of the labor force cannot be solved by simple investment, as mentioned above. This study intends to investigate how smartwork will improve productivity in the ageing society.

3 Smartwork: Advanced Telework

Smartwork is a form of advanced telework, although there is no standard definition as yet. Conceptually, smartwork can be defined as flexible working in which people can work with anybody using ICT on the network, without being constrained by time or location. By embodying collective intelligence and encouraging the economic contribution of both the female and disabled labor force, smartwork can be utilized as an alternative means of coping with the problem of the declining economic population due to the low birthrate and ageing, and thus has the potential to improve the quality of life of each citizen. Smartwork breaks away from the typical working pattern of working in a designated location during a designated period, and has the work follow the worker. While telework is the concept of pursuing trust-based management and efficiency through ICT-based remote working, smartwork is recognized as a concept that adds the human-oriented system or culture. In that respect, smartwork is a concept that focuses more on work that is flexible in terms of time and location, and which can be considered as pursuing the improvement of labor efficiency based on mutual trust by integrating and utilizing various types of information and knowledge.

Smartwork can be categorized as telecommuting, mobile office, and workplace working based on flexibility of time and location. Conceptually, it is often used analogously with teleworking, telecommuting and flexible working. Furthermore, it includes a business solution that enables remote collaboration such as “Telepresence” and Web conferencing, as well as mobile office solutions using smart phones; tablet PCs, laptops, and so on.

Workplace, a type of smartwork, means working in a facility that has an ICT-based teleworking system near the residence. It features a composite space of the same level as a conventional work space, thus providing the office space needed for knowledge work so as to improve the concentration of the workers and make it easier to manage work hours, and thereby solving the problem of telecommuting. In the Netherlands, a type of workplace called DoubleU has been installed in 99 locations in major urban areas as of 2010 [10].

Many factors have been presented as the effects of smartwork. A change in methods of working based on cutting-edge ICT increases the satisfaction of workers and is an important strategy for reducing CO₂ emissions and solving population concentration at the national level. From the productivity perspective, smartwork has very high potential because it can not only increase productivity by improving the

work efficiency of the workers, but can also greatly increase the productivity of the aged labor force to solve the decline in productivity due to the reduction of the labor force. The next chapter will discuss policy solutions for the successful introduction and establishment of smartwork.

4 Policies for the Success of Smartwork

There are several issues to review in order to successfully introduce smartwork through ICT investment. First, smartwork has a similar problem as telework. The connectivity problem arising from the distributed work arrangement can also occur in the smartwork environment. And there is the emotional issue of workers as they work in a place that is not a conventional work site. Teleworkers are known to have greater feelings of loneliness, irritation, concern, and guilt than those working in a conventional office. Since smartwork requires a stronger continuity of the working environment than telework, an enterprise will always have to worry about security issues. The information security problem that can occur when utilizing the various information and communication services such as the cloud service and the social network service, in addition to the security problem of the shared workspace, is an important problem that smartwork must resolve.

To solve such a problem, the following policies should be implemented. First, a work management policy appropriate for smartwork is needed. Since smartwork is not just a teleworking type, the working form can be flexibly adopted according to the business type. Therefore, the means of managing a personalized operation instead of a uniform working type is needed. This does not just mean securing the right framework but also assuring the right of the individual worker to make the choice. Second, a performance-oriented evaluation policy that can break away from the face-to-face culture must be adopted. Smartwork cannot succeed if the worker's performance appraisal depends on non-performance-related factors such as one's relationship with the boss. Although many measures to mitigate such a problem have been introduced through trial and error in the continued drive to promote telework, smartwork must improve the performance-oriented appraisal even more. Lastly, the promotion of smartwork with consideration of the public interest and universality is needed. The problem of the public interest and universality has been continuously raised as telework has expanded its base. A flexible working policy centered on public agencies may not induce sympathy from most workers, and driving forward individual policies without a national standard could lead to confusion. Particularly since the security and information protection problem mentioned above is the most critical issue for enterprises, there will be many difficulties in expanding the basis of smartwork unless there is a universal security guideline and standard.

5 Conclusions

This paper discusses the smartwork policy required to improve productivity in the era of the ageing society. ICT investment to increase productivity and development of the

smartwork environment are accepted as effective investment targets. Although smartwork is still in its infancy, it is expected that it will proliferate along with the national effort to expand its base and show the visible success of smartwork mostly in the ICT industry.

Smartwork can be considered as an advanced form of telework that enables remote collaboration based on ICT. As mentioned above, there are some policy problems that must be resolved in order to adopt smartwork successfully. Moreover, there is no clear standard to clearly conclude that smartwork will help increase productivity just by solving such problems. As such, it is probably premature to drive ahead with smartwork and establish the criteria required to determine its success in improving productivity, as it is not yet clear to what extent and when smartwork can be adopted in different industries. However, the experience accumulated by public agencies and commercial enterprises in adopting telework will be greatly helpful in solving the potential problems and should eventually contribute to effectively improving productivity.

Future studies will include a quantitative study of low productivity among the aged work force in a bid to analyze how the aged work force influences decreases in productivity, and an empirical study of the ways in which smartwork can contribute to the improvement of the productivity of the aged work force.

References

1. United Nations, World Population Prospects: The 2010 Revision (2011)
2. Ministry of Health, Labour and Welfare, Special Report for Dynamics of Population Statistics (2002), <http://www.mhlw.go.jp/>
3. OECD, Unit Labour Costs – Annual Indicators: Labour Productivity per Unit Labour Input (2011), <http://stats.oecd.org/Index.aspx>
4. <http://kostat.go.kr/> (2011) (in Korean)
5. Berndt, E.R., Morrison, C.J.: High Tech Capital Formation and Economic Performance in U.S. Manufacturing Industries. *An Exploratory Analysis Journal of Econometrics* 65(1), 9–43 (1995)
6. Brynjolfsson, Hitt: Information Technology and Its impact on Firm-level Productivity: Evidence from Government and Private Data Source 1977-1993. *Canadian Journal of Economics* 32(2), 335–362 (1999)
7. Sircar, S., Turnbow, J.L., Bordoloi, B.: A framework for assessing the relationship between information technology investments and firm performance. *J. Manage. Inf. Syst.* 16(4), 69–97 (2000)
8. Nordhaus, W.D.: Productivity Growth and the New Economy, NBER Working Paper No. 8096 (2001)
9. Becchetti, Bedoya, Paganetto: ICT Investment, Productivity and Efficiency: Evidence at Firm Level Using a Stochastic Frontier Approach. *Journal of Productivity Analysis* 20 (2003)
10. http://www.connectedurbandevlopment.org/connected_and_sustainable_work/double_u_smart_work_network

Erratum: Idea Sketch of Application Service of e-Business Card and NFC Based on Mobile

Sun Kuh Noh¹, Byung Rae Cha², Dong You Choi³, Jung-Jin Park³,
Guen Sik Kim³, and Yoo Kang Ji⁴

¹ Dept. of Radio Mobile Communications Engineering, Honam University,
#330 Eodeungno, Gwangsan-gu, Gwangju, 506-714, Korea
nsk7070@honam.ac.kr

² SCENT Center, GIST, GwangJu, Korea
brcha@nm.gist.ac.kr

³ Dept. of Information & Communications Engineering, Chosun University
#375 Seosuk-dong, Dong-gu, Gwangju, 501-759, Korea

⁴ Dept. of Information & Communications Engineering, DongShin University
#252 Deaho-dong, Naju, Jeonnam, 520-220, Korea
neobacje@gmail.com

T.-h. Kim et al. (Eds.): UNESST 2011, CCIS 264, pp. 66–74, 2011.
© Springer-Verlag Berlin Heidelberg 2011

DOI 10.1007/978-3-642-27210-3_46

In the original version, the name of the first author is incorrect. Instead of “Sun Kuh Noh” it should be read as “Sun Kuk Noh”.

The original online version for this chapter can be found at
http://dx.doi.org/10.1007/978-3-642-27210-3_8

Author Index

- Abawajy, Jemal 83
Ahmad, Masitah 83

Bai, Lin 130

Cha, Byung Rae 66, E1
Chae, Seong Wook 168, 177
Chan, Veng-Ian 31
Chantrapornchai, Chantana 106, 114,
122, 153
Che, Zhongzhi 93, 100
Chen, Chao-Chun 75
Chen, Kai 93
Chen, Rain 75
Cho, Min-Hee 258
Choi, Dong You 66, E1
Choi, Do Young 207, 217
Choi, Sung-Pil 233, 269
Choi, Yun-Soo 233, 269
Chun, Hong-Woo 233, 269

Dai, Ting-Ting 75
Dolwithayakul, Banpot 122

Esmaeilzadeh, Pouyan 17

Guan, Wenyang 130

Hahn, Min Hee 207, 217
He, Jianhua 93, 100, 130
Hwang, Mi-Nyeong 258
Hwang, Myung-Gwon 258, 287
Hwang, Sung-Hyun 316

Jeong, Chang-Hoo 233, 269
Jeong, Do-Heon 258, 287
Jeong, Soon hwan 10
Ji, Yoo Kang 66, E1
Jo, Hyun-Woong 332
Jo, Nam Yong 187, 197
Jung, Hanmin 245, 263, 281, 294, 301
Jung, Su-Hyun 306

Kanchan, K. 114
Kang, In-Su 239, 251, 274

Kang, Sin-Jae 251
Kang, Sung-Lyong 306
Khan, W.A. 1
Khattak, A.M. 1
Kim, Guen Sik 66, E1
Kim, Jinhyung 263, 301
Kim, Kwang-Young 227
Kim, Pyung 245, 263, 281
Kim, Sang-Wook 332
Kim, Sung-Hong 306
Kim, Taehong 245, 263
Kim, Tai-hoon 83
Kong, Hee-kyung 341
Koo, Hee Kwan 263, 281
Kumar, Naresh 17
Kwak, Yoonsik 144
Kwak, Youngsik 136, 144
Kweon, Soo Cheon 324

Lee, Dae Sung 187, 197
Lee, Hyeongjik 324
Lee, Jinhee 263, 301
Lee, Kun Chang 168, 177, 187, 197,
207, 217
Lee, Mikyoung 258, 263, 294, 301
Lee, Min-Ho 227, 230
Lee, Seungwoo 245, 263, 281, 294, 301
Lee, S.Y. 1
Lee, Won Bin 324
Lee, Won-Goo 227
Lee, Y.K. 1
Lee, Youngkon 42, 50, 58

Ma, Chao 100
Mantrawong, O. 106
Marigomen, Ralph 274
Meurant, Robert C. 160

Na, Seung-Hoon 239, 251
Nam, YoonJung 136
Nam, Youngjin 10
Nezakhati, Hossein 17
Noh, Sun Kuh 66, E1

Park, Jung-Jin 66, E1

Park, Young kyun 10

Pervez, Z. 1

Ryou, Jae-Cheol 230

Sambasivan, Murali 17

Sathapanawat, T. 153

Seo, Dongmin 263, 281

Seo, Young Wook 168, 177

Shin, Sung-Ho 227

Si, Yain-Whar 31

Song, Sa-Kwang 233, 269

Song, Seokil 144

Sudta, P. 114

Sung, Kihoon 341

Sung, Won-Kyung 227, 230, 233, 245,
263, 269, 281, 287, 294

Tang, Zuoyin 93, 100, 130

Yeon, Seung-Jun 316

Yoon, Hwa-Mook 227, 230

Yooyeunyong, Suabsakul 122

Zhou, Yi 93, 100, 130