Junzo Watada
Toyohide Watanabe
Gloria Phillips-Wren
Robert J. Howlett
Lakhmi C. Jain
Editors

SMART INNOVATION, SYSTEMS AND TECHNOLOGIES ■ 15

Intelligent Decision Technologies

Proceedings of the 4th International Conference on Intelligent Decision Technologies (IDT'2012) – Volume 1

Springer

Smart Innovation, Systems and Technologies 15

Editors-in-Chief

Prof. Robert J. Howlett
KES International
PO Box 2115
Shoreham-by-sea
BN43 9AF
UK
E-mail: rjhowlett@kesinternational.org

Prof. Lakhmi C. Jain
School of Electrical and Information Engineering
University of South Australia
Adelaide
South Australia SA 5095
Australia
E-mail: Lakhmi.jain@unisa.edu.au

For further volumes:
http://www.springer.com/series/8767

Junzo Watada, Toyohide Watanabe,
Gloria Phillips-Wren, Robert J. Howlett,
and Lakhmi C. Jain (Eds.)

Intelligent Decision Technologies

Proceedings of the 4th International Conference on Intelligent Decision Technologies (IDT'2012) – Volume 1

Springer

Editors
Prof. Junzo Watada
Waseda University
Kitakyushu
Japan

Prof. Toyohide Watanabe
Nagoya University
Japan

Prof. Gloria Phillips-Wren
Loyola University Maryland
Baltimore, MD
USA

Prof. Robert J. Howlett
KES International
Shoreham-by-sea
United Kingdom

Prof. Lakhmi C. Jain
University of South Australia
Adelaide
Australia

ISSN 2190-3018
e-ISSN 2190-3026
ISBN 978-3-642-29976-6
e-ISBN 978-3-642-29977-3
DOI 10.1007/978-3-642-29977-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012937732

Printed on acid-free paper

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

Preface

The Intelligent Decision Technologies (IDT) International Conference encourages an interchange of research on intelligent systems and intelligent technologies that enhance or improve decision making. The conference is organized by KES International, a research community consisting of several thousand research scientists. The focus of IDT is interdisciplinary and includes research on all aspects of intelligent decision technologies, from fundamental development to real applications.

Advances in Artificial Intelligence (AI) and computing environments that can deliver intelligent technologies effectively have enabled an explosion in intelligent applications. IDT have the potential to expand their support of decision making in such areas as finance, accounting, marketing, healthcare, medical and diagnostic systems, military decisions, production and operation, networks, traffic management, crisis response, human-machine interfaces, financial and stock market monitoring and prediction, and robotics. Intelligent decision systems implement advances in intelligent agents, fuzzy logic, multi-agent systems, artificial neural networks, and genetic algorithms, among others. Emerging areas of active research include virtual decision environments, social networking, 3D human-machine interfaces, cognitive interfaces, collaborative systems, intelligent web mining, e-commerce, e-learning, e-business, bioinformatics, evolvable systems, virtual humans, and designer drugs.

In this volume we publish research from the Fourth KES International Symposium on Intelligent Decision Technologies (KES IDT'12), hosted and organized by researchers in Japan. The conference was held in Gifu City located in the center of Japan. Gifu City is known for its traditions and rich history, including its 1300-year-old tradition of cormorant fishing on the Nagara River and Gifu Castle. This book contains chapters based on papers selected from a large number of submissions for consideration for the symposium from the international community. Each paper was double-blind, peer-reviewed by at least two independent referees. The best papers were accepted based on recommendations of the reviewers and after required revisions had been undertaken by the authors. The final publication represents the current leading thought in intelligent decision technologies.

We wish to express our sincere gratitude to the plenary speakers, invited session chairs, delegates from all over the world, authors and reviewers for their outstanding contributions. We express our sincere thanks to the Japanese organizers and Gifu City for their sponsorship and support of the symposium. We thank the International Programme Committee for their support and assistance. We would like to thank Peter Cushion of KES International for his help with organizational issues. Also our appreciation is extended to their contribution provided by the KES International Secretariat Team. We thank the editorial team of Springer-Verlag for their support in production of this volume. We sincerely thank the Local Organizing Committee, especially Professors K. Asakura in Daido University and T. Kojiri in Kansai University for their editing conference record and program, and their collaborators for their invaluable contributions.

We hope and believe that this volume will contribute to ideas for novel research and advancement in intelligent decision technologies for researchers, practitioners, professors and research students who are interested in knowledge-based and intelligent engineering systems. At the end we will express thanks for Dr. Jerzy Michnik, University of Economics in Katowice, Poland, for his help on Latex.

Gifu, Japan
May 22–25, 2012

Junzo Watada
Toyohide Watanabe
Gloria Phillips-Wren
Robert J. Howlet
Lakhmi C. Jain

Editors

Professor Junzo Watada received his B.Sc. and M.Sc. degrees in electrical engineering from Osaka City University, and Dr. of Eng. Degree from Osaka Prefecture University, Japan. He is currently a Professor of Knowledge Engineering, Soft Computing and Management Engineering at the Graduate School of Information, Production & Systems, Waseda University, after a professor of Human Informatics and Knowledge Engineering at the Osaka Institute of Technology, Japan and was with Ryukoku University, Kyoto for about 10 years each. Before moving to Academia, he was with Fujitsu Co. Ltd. for about 10 years. His research interests includes decision making tehnologies and management of technology. Dr. Watada is currently the President of International Society of Management Engineers. He was a co-chair of KES-IDT2011 at Pireus, Greek and plays an active role in editing International Journal of Intelligent Decision Technology (KES official journal) as a co-editor-in-chief.

Toyohide Watanabe has received B.S., M.E. and Dr.Eng. from Kyoto University, in 1972, 1974 and 1985, respectively. Since 1975, he has worked in Kyoto University and Nagoya University: first as a research associate in Data Processing Center, Kyoto University, from 1975 to 1987; second as an associate professor Faculty of Engineering, Nagoya University, from 1987 to 1994; third as a full professor in the same faculty, from 1994 to 1997; fourth as a professor in Graduate School of Engineering, Nagoya University

from 1997 to 2003; and finally as a professor in Department of Systems and Social Informatics, Graduate School of Information Science, Nagoya University, since 2003. Additionally, from 2004 to 2008 he was the head director in Information Technology Center, Nagoya University. The current topics of his research interests are: Intelligent Tutoring System, Computer-supported collaborative learning, Knowledge management, Intelligent activity-support, etc. He is a member of ACM, IEEE-CS, AAAI, AACE, KES International, IEICE of Japan, IPS of Japan, IEE of Japan, Japan of SAI, Japan of SSST, JSiSE, etc. Also, currently he is an Editor-in-Chief of the International Journal of Knowledge and Web Intelligence. Moreover, he is a fellow on IEICE of Japan since 2004.

Dr. Gloria Phillips-Wren is Professor and Chair of Information Systems and Operations Management at Loyola University Maryland, and Academic Director of Executive Programs. She is co-founder and co-editor-in-chief of Intelligent Decision Technologies International Journal (IDT), Vice Chair and Chair-elect of SIGDSS under the auspices of the Association for Information Systems (AIS), Secretary of IFIP WG8.3 DSS, and leader of the focus group in Intelligent Decision Technologies (IDT) for KES International. She is a co-organizer of the IDT Conference Series under the auspices of KES International. She received a PhD from the University of Maryland Baltimore County and holds MS and MBA degrees. Her research interests and publications are in intelligent decision support systems, intelligent agents, decision making, analytics, business intelligence, data mining, and emerging technologies such as social media. She has published in the area of decision making and support in academic journals including the European Journal of Operational Research, Omega, Expert Systems with Applications, and the Journal of Network and Computer Applications. Her most recent book (co-edited) was published in 2011 and is entitled Intelligent Decision Technologies.

Robert Howlett is the Executive Chair and Chief Executive of the KES (Knowledge-Based and Intelligent Engineering Systems) International. KES is an associaton dedicated to knowledge exchange and the dissemination of research results in the areas of intelligent systems, sustainability in energy and buildings and innovation through knowledge transfer. He is a Director of the Institute of Knowledge Transfer, the body for knowledge professionals and researchers working in innovation, knowledge transfer and enterprise. He holds a Visiting Professorship in Enterprise at Bournemouth

University. He holds a PhD in Intelligent Systems from the University of Brighton, an MPhil from Sussex University and a BSc(Hons) from Portsmouth Polytechnic. He is a Chartered Engineer and a Chartered Information Technology Practitioner. He was at the University of Brighton for over 20 years where he headed a research group in intelligent systems and the was Director of the Knowledge Transfer Partnerships (KTP) Centre. For a number of years he chaired the KTP national Forum, representing KTP centres from all universities in the country, which he formed. He has been involved with innovation knowledge transfer, as a practitioner, facilitator and manager, for over 15 years. He has personally supervised many knowledge transfer projects and written and mentored many more.

Professor Lakhmi Jain is a Director/Founder of the Knowledge-Based Intelligent Engineering Systems Centre. His interests focus on the novel techniques such as knowledge-based intelligent machines, virtual systems, intelligent agents, and the application of these techniques in areas such as engineering, science, aviation, healthcare, defence and so on.

Contents

Part I: Methodologies of Intelligent Decision Technology

Part II: Applications of Intelligent Decision Technology

Part I
Methodologies of Intelligent Decision Technology

A Compromise Decision-Making Model to Recover Emergency Logistics Network

Yiping Jiang and Lindu Zhao

Abstract. Quick recovery of emergency service facilities (ESFs) in the aftermath of large-scale disasters has emerged as a hot topic in the field of emergency logistics. This paper focuses on the ESFs recovery problem under the constraints of scare emergency resource and recovery time. In this study, a compromise programming model is proposed as an integrated decision support tool to obtain an optimal compromise solution with regard to two objectives: minimize the consumption of recovery resources and maximize the resilience capacity through selecting different recovery strategies. Then a genetic algorithm is proposed to solve the developed mathematical model, and a numerical example is followed to illustrate the effectiveness and usefulness of proposed model.

1 Introduction

Emergency logistics network as a key guarantee to respond various disasters [1], it contains several physical emergency service facilities (ESFs). Most ESFs play a central role in emergency aid, crowd evacuation and rescue materials distribution, etc. But those spatially distributed ESFs are also vulnerable to various large-scale disasters, such as earthquakes, tsunamis and hurricanes. Such disasters can damage those ESFs and affect their capacity to satisfy the needs for which they were developed, and then halt the progression of post-disaster rescue and repair activities. Hence, how to recover the damaged ESFs so as to enhance resilience capability for the emergency logistics network is a great challenge in the field of emergency management [2].

Resilience as a metric for emergency recovery problem, it is defined as the capacity of physical and human systems recovered from extreme events [3]. Many scholars in the filed of emergency management have increasingly paid attention to

Yiping Jiang · Lindu Zhao
Institute of Systems Engineering, Southeast University, Nanjing, China, 210096
e-mail: ypjiang@seu.edu.cn, ldzhao@seu.edu.cn

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 3–12.
springerlink.com

resilience study [2, 3, 4, 5]. Although the topic of resilience has emerged a few years ago, its study related to emergency management has only been widely recognized recent years [6], in general, whose basic idea is considered as a capability to restore to the normal states [7]. With more and more researchers' participant, many analysis methods are widely applied into post-disaster recovery analysis and modeling, such as statistical recovery curves analysis [8], resource constraints approach [9], and optimization programming model [10, 11]. Generally, most of current researches were qualitative analysis, and some few quantitative studies also were single objective optimization; on the other hand, some studies only investigated the resilience definition or mechanisms, but how to enhance the resilience capability is still a problem. In fact, available emergency resources and resilience capability are the key criteria to assess the recovery strategies. In some sense, resilience capability is strictly depending on the resources employed, the more resource redundance, the higher resilience improved [12]. But, in the context of large-scale disasters, emergency resource is becoming strictly scarce when confronting the demand surge from mass disaster-affected regions. Hence, economically use scare recovery resource and greatly improve the resilience capability are conflicted each other for recovery decision-making, and should be balanced together.

Compromise Programming (CP) is firstly proposed with many theoretical extensions and with applications in several fields in 1970s [13, 14, 15, 16]. It is a useful technique to determine a compromise set that is nearest with respect to an ideal and infeasible point, for which all the criteria are balanced and optimized in Pareto points. In this paper, we restrict our attention to tradeoff the objectives between consumption of emergency resource and resilience capability for emergency logistics network recovery problem in the context of large-scale disaster. The purpose of this study is to propose a CP-based model to support the emergency recovery decision making in the context of large-scale disaster.

The remainder of the paper is organized as follows. Section 2 describes a CP-based programming model. Section 3 provides an efficiency solution algorithm to solve the model proposed. A numerical example and some computational results are discussed in Section 4. Finally, we give conclusions and directions for future research in Section 5.

2 Model Formulations

In this section, we propose a CP-based programming model to assist the emergency recovery decision-making. At first, we formulate the different independent optimizations, i.e., minimize the consumption of recovery resources and maximize the resilience capacity. Then, we propose a CP-based model to tradeoff the two objectives. Basic assumptions are described as follows:

Firstly, we assume the emergency resource being used in every recovery strategy is a general concept, and do not divide into various categories in the model.

Secondly, each recovery strategy is assumed as an integrated process or plan, and do not consider the specific sub-process or sub-plan in this paper.

Thirdly, the proposed model is fit for real-time decision, and we only consider the decision circumstance at time *t*.

(1) Notations

According to the definition in [4], the loss of resilience is the size of the function degradation of facilities over time from disaster happened time t_0 to completely recovery time t_1, $\int_{t_0}^{t_1}[100-Q(t)]dt$. Inspired by this work, we note *100-Q(t)* as the loss function of the emergency logistics network at time *t*. Here, we define $Q_i(t)$ as the function for each emergency service facility at time *t*; *Q(t)* is the function of an emergency logistics network composed of some ESFs at time *t*, which can range from 0% to 100%, where 100% means no damage of ESFs and 0% means complete damage of ESFs. Other parameters and variables are given as follows.

K: number of ESFs in an emergency logistics network;
J: set of recovery strategies;
Z_r: total amount of available emergency recovery resources;
Q: minimum desired recovery capability for an emergency logistics network composed of various ESFs.
T: time limitation for recovering all ESFs;
E_i: damaged area of ESFs *i*;
R_{ij}: resilience capability for ESFs *i* adopted by recovery strategy *j*;
C_{ij}: required emergency recovery resource of per unit area for ESFs *i* by using recovery strategy *j*;
T_{ij}: recovery time for per unit area of ESFs *i* by using recovery strategy *j*;
X_{ij}: a binary parameter (0, 1) whether recovery strategy *j* provides recovery capability for ESFs *i*;

(2) Model I: Independent-Objective-Based Models

First of all, we propose the model to minimize the total recovery resource consumption.

$$f_1(x)=\min\sum_{i\in K}\sum_{j\in J}C_{ij}E_iX_{ij} \tag{1}$$

Subject to:

$$\sum_i\sum_j C_{ij}E_iX_{ij}-Z_r\le 0,\quad i\in K, j\in J \tag{2}$$

$$\sum_i\sum_j R_{ij}E_iX_{ij}\Big/\sum_i E_i-Q\ge 0,\quad i\in K, j\in J \tag{3}$$

$$\sum_i\sum_j T_{ij}E_iX_{ij}-T\le 0,\quad i\in K, j\in J \tag{4}$$

$$\sum_j X_{ij}=1,\quad \forall i\in K \tag{5}$$

Eq. (1) is to minimize total recovery resource consumption. Constraint (2) is a resource constraint. Constraint (3) ensures the resilience capability exceed the minimum specified value Q. Constraint (4) restricts that the total recovery time. Constraint (5) states that at most one recovery strategy can be selected.

Then, we formulate the following model to maximize the resilience capability through maximizing the total recovery capability of an emergency logistics network composed of the ESFs.

$$f_2(x) = \max \left\{ \frac{\sum_{i \in K} \sum_{j \in J} R_{ij} E_i X_{ij}}{\sum_{i \in K} E_i} \right\} \quad \text{st. (2), (3), (4), (5)} \tag{6}$$

(3) Model II: CP-Based Model

Based on the above discussion, here we propose a CP-based model to achieve the "best compromise" solution between different objectives. The main idea of CP technique is the competitive nature of mainly conflicting objectives that compete for the same limited resources [15]. Thus, according to the principle of distances metric "L_p", the proposed model in this section is defined as follows:

$$\min L_p(\overline{\lambda}, x) = \left[\sum_i \lambda_i^p \cdot \left| \frac{f_i^* - f_i(x)}{f_i^* - f_{i^*}} \right|^p \right]^{\frac{1}{p}}, \quad i = 1, 2 \quad \text{st. (2), (3), (4), (5)} \tag{7}$$

Where $L_p(\overline{\lambda}, x)$ represents the performance index measure the optimization of the multi-objective function; p denotes the distance metric enabling to implement various distance measures into the model, which is the integer parameter with values in $[1, \infty]$; $\overline{\lambda}$ is the vector of weights of representing the relative importance of the discrepancy between each relevant objective and its corresponding ideal value, and $\sum_{i=1}^{2} \lambda_i = 1, \lambda_i \geq 0$; f_i^* and f_{i^*} are the ideal value and anti-ideal value for the i^{th} objective in the independent optimization model, respectively.

3 GA-Based Solution Method

Our model is a mixed integer combinatorial optimization problem, which is NP hard to solve in practice. Since genetic algorithm (GA) is a common algorithm to optimize various objective functions [17, 18], we here adopt it to design the solution procedures.

(1) Chromosome Coding

Each chromosome i contains m genes, where m is the number of ESFs. The position of each gene in a chromosome denotes the corresponding recovery strategy, and has K values, where K is the number of recovery strategies. For example, if there are 8 ESFs and 4 recovery strategies, a chromosome coding of 41232341 indicates that the 1st UESF is adopted recovery strategy 4, and so on.

(2) Selection Operator

Step One: Given the parameter a within (0, 1), the evaluation function based on the orders is defined as $eval(h_i, f_{h_i}) = a(1-a)^{i-1}, i = 1, 2, \cdots, N$,

Step Two: For each chromosome, calculate the cumulative probability.

$$p_0 = 0, p_i = \sum_{j=1}^{i} eval(h_j, f_{h_j}), i = 1, 2, \cdots, N.$$

Step Three: Generate a real number r belongs to $(0, p_n)$ randomly.

Step Four: If r locates in $(p_{i-1}, p_i]$, then select the chromosome h_i.

Step Five: Repeat the Step 3 and Step 4..

(3) Crossover Operator

Step One: Randomly select the crossover partner and exchange the code on them with probability p_c.

Step Two: Checking up whether constraints (2), (3) and (4) are satisfied.

Step Three: Repeat Steps 1 and 2.

(4) Mutation Operator

Step One: Randomly selected the mutation chromosome and vary in certain places with probability p_m.

Step Two: Check the chromosomes through constraints (2), (3) and (4). If the new offspring is illegal, restore the parent chromosomes before mutation operation, and restart a new mutation process once time until the offspring is satisfied the constraints (2), (3) and (4).

4 A Numerical Example

A hypothetical emergency logistics network composed of twelve ESFs ($m=12$) is proposed to verify the post-disaster recovery model. We assume ESFs' function (i.e. $Q_i(t)$) is described on a discrete and ordinal scale of 1, 2, 3, 4, and 5 to represent the states of "Serious", "Bad", "Moderate", "Good", and "Very Good", respectively. The indicator of resilience capability given by 1, 2, 3, 4, and 5 is used to describe the extent of damaged area being less than or equal to100%, 60%, 30%, 10% and 5%, respectively. Correspondingly, the time relevant to safety including "especial", "considerable", "major", "restricted" and "minor" is also represented by 1, 2, 3, 4, and 5, respectively. Based on the rating sequences, an appropriate recovery strategy for ESFs can be determined amongst "integrated recover", "corrective recover", "replaceable recover", "preventive recover" and "routine maintenance", that's to say, $K=5$.

To verify the proposed formulations, we generate the some data by using uniform distribution in MATLAB7.1 (see Table 1). The extent of damaged area (E_i) varies between 100 m^2 and 130 m^2; the function ratings of ESFs at time t, $Q_i(t)$, vary between 1.0 and 5.0. Besides, the value of resilience capability (R_{ij}) also varies between 1.0 and 5.0 for recovery strategies 1~ 4, and the fifth recovery strategy is set as 5. The consumption of resource (C_{ij}) varies within 0.8~3.0, 3.0~6.0, 4.0~10.8, 9.0~12.8 for the second, third and fourth strategy, and the fifth strategy is set as zero score due to no emergency resource is required; recovery time (T_{ij}) for each strategy is similar with the value of resilience capability. Other parameter settings are also given in Table 2.

Table 1 Parameter settings for ESFs and recovery strategies

i	E_i	$Q_i(t)$	R_{ij}					C_{ij}					T_{ij}				
			M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5
1	100.34	2.0	1.8	2.5	3.6	4.3	5.0	0	2.58	5.25	10.01	12.49	0	0.85	0.55	0.33	0.24
2	115.46	1.2	1.0	1.7	3.0	4.1	5.0	0	2.90	6.41	10.56	12.70	0	0.84	0.65	0.42	0.25
3	110.99	2.4	2.1	2.9	3.2	4.2	5.0	0	2.11	3.23	9.04	11.96	0	0.78	0.51	0.34	0.19
4	128.41	2.9	2.5	3.6	4.3	4.8	5.0	0	1.97	3.04	9.04	11.33	0	0.87	0.68	0.49	0.26
5	108.96	2.8	2.3	3.5	4.2	4.8	5.0	0	2.10	3.56	9.30	12.05	0	0.87	0.51	0.49	0.19
6	128.13	4.3	4.1	4.3	4.5	5.0	5.0	0	1.05	3.81	8.01	9.98	0	0.89	0.57	0.33	0.24
7	117.24	3.2	2.9	3.7	4.6	4.9	5.0	0	1.91	3.79	8.73	11.03	0	0.86	0.65	0.35	0.11
8	111.44	3.4	3.2	3.9	4.6	5.0	5.0	0	1.94	5.42	8.29	10.87	0	0.72	0.57	0.42	0.15
9	100.22	1.1	0.5	1.9	3.0	4.0	5.0	0	2.64	4.09	7.91	12.76	0	0.76	0.54	0.36	0.13
10	119.66	5.0	4.7	5.0	5.0	5.0	5.0	0	0.89	2.80	4.57	9.60	0	0.82	0.64	0.49	0.16
11	127.69	1.9	1.8	2.5	3.4	4.5	5.0	0	2.23	3.06	7.58	12.18	0	0.88	0.51	0.43	0.27
12	123.06	3.0	2.5	3.5	4.3	4.9	5.0	0	2.58	5.99	8.85	10.93	0	0.71	0.68	0.33	0.19

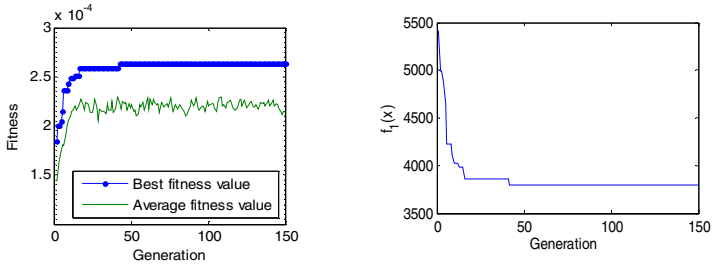
Table 2 Parameter settings of the constraints and the compromise model

Parameter	Z_r	Q	T	λ_1	λ_2	p
Value	9000	3.8	600	0.6	0.4	2

(1)Optimization Results of Model I

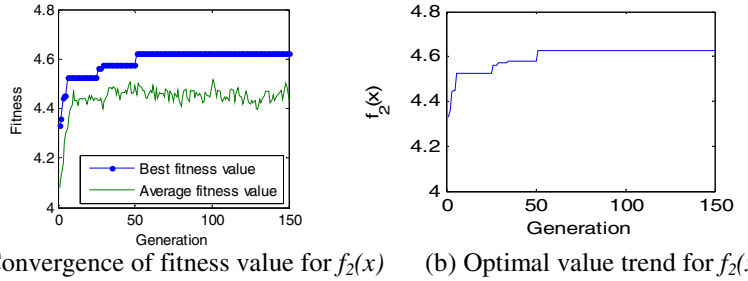
Initial values of parameters in applying GA are given as follows: $N=50$, $p_c=0.6$, $p_m=0.05$, iteration number set as 150, and suppose their values are unchanged during the iteration process. After iterating 100 times, the optimal solution for each objective is obtained, depicted in Fig.4. In Fig.4 (a), the dotted line states that the average fitness value of $f_1(x)$ is increasing with the iterative generations, and the solid line shows the diversity of populations during iterations. On the other

hand, the curve demonstrated in Fig.4 (b), it reflects the varying trend of optimal solution, $f_1(x)=3791.6$. Similarly, the results shown in Fig.5 can be analyzed from the same perspectives, and the optimal solution for $f_2(x)$, and $f_2(x)=4.63$.



(a) Convergence of fitness value for $f_1(x)$ (b) Optimal value trend for $f_1(x)$

Fig. 1 Convergence of fitness value, and optimal value trend for $f_1(x)$



(a) Convergence of fitness value for $f_2(x)$ (b) Optimal value trend for $f_2(x)$

Fig. 2 Convergence of fitness value, and optimal value trend for $f_2(x)$

In order to compare the conflict between resilience and efficiency during emergency recovery process clearly, the ideal (best) and anti-ideal (worst) values for each of the objectives is compared in Table 3.

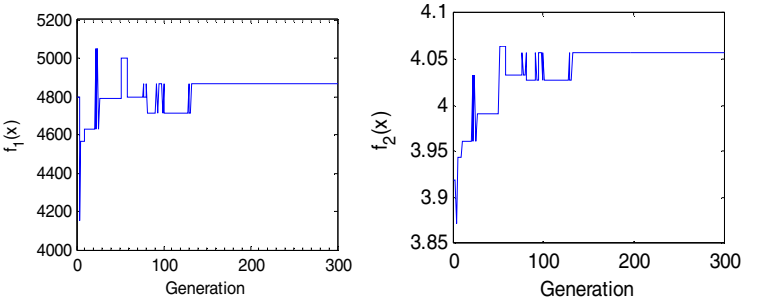
Table 3 The ideal and anti-ideal values for each objective

Objectives	Ideal	Anti-ideal
$f_1(x)$	3791.6	8922.46
$f_2(x)$	4.63	3.82

(2)Optimization Results of Model II

We note that the basic parameters for GA as follows: $N=50$, $p_c=0.6$, $p_m=0.05$, iteration number set as 300, and assume their value are not changed during the iteration. We again iterate the problem 300 generations, and then the compromise process for $f_1(x)$ and $f_2(x)$ are shown in Fig.3. Moreover, the changing trend of the

compromise solution between these two objectives is also depicted in Fig.4. At this time, the optimal Pareto solutions are $f_1(x)=4868.22$, $f_2(x)=4.06$.



(a) Compromise solution of $f_1(x)$ (b) Compromise solution of $f_2(x)$

Fig. 3 Tradeoff process between $f_1(x)$ and $f_2(x)$

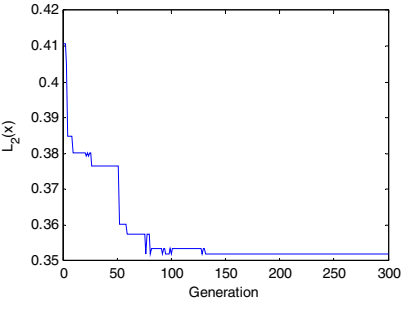


Fig. 4 Optimal value trend of distance metrics L_2

Moreover, the comparison of optimization results between Model I and Model II is given in Table 4.

Table 4 Comparison of optimal results

The objectives	Model I	Model II
$f_1(x)$	3791.6	4868.22
$f_2(x)$	4.63	4.06

(3)Key Parameter Analysis

It is quite evident to see that the distance metric, p, is a key parameter to the decision of the compromise solution. In this section, a sensitivity analysis of distance metric p on the balance of $f_1(x)$ and $f_2(x)$ during post-disaster recovery is conducted. Holding all the other parameters fixed, we have obtained the compromise decisions and results in Table 5.

Table 5 Key parameter analysis

The objectives	$p=1$	$p=2$	$p=\infty$
$f_1(x)$	6844.83	4868.22	4652.15
$f_2(x)$	3.81	4.06	4.13

As indicated in Table 5, $f_2(x)$ raises with the increase of distance metric, because the solution trades the emergency resource and recovery time for resilience capability when $p \to \infty$, while the $f_1(x)$ is gradually decreased. Moreover, the solution corresponding to $p=2$ strikes a better balance between the two objectives than the solution corresponding to $p=1$ and $p=\infty$, in the sense of leveling the percentage achievement of the different objectives from their respective ideal values and can thus be more appealing to the planner. Hence, the solution when $p=2$ can be thought as the best compromise solution.

As the above analysis demonstrates, the distance metric plays an important role in compromise decisions between $f_1(x)$ and $f_2(x)$ during post-disaster recovery decision making. Unfortunately, predicting an accurate distance metric p for a compromise solution is difficult and is usually done empirically. As the accuracy of the distance metric p is vital to the success of decision-making for recovering emergency logistics network, a great deal of effort needs to be devoted to scientifically estimating the distance metric of different recovery scenarios.

5 Conclusions

In this study, we develop a CP-based decision-making model to recover emergency logistics network in the context of large-scale disaster. The decision-making model seeks the acceptable compromise solution between minimization of recovery resource and maximization of resilience capability with the constraints of scarce recovery resource and time limitation for emergency recovery activities. Moreover, we also design a GA-based solution procedure to solve the combinatorial optimization problem contained in the proposed model. Then a numerical example is given to illustrate the applicability of our decision-making model, and the results and analysis verify our model is useful in practice.

Moreover, our research can be extended in several directions by relaxing the assumptions in this research. First of all, we only consider the general recovery resource. Actually, the recovery resource is classified into different kinds in practice. When different kinds of resources are considered, the decisions on selecting recovery strategies will be more complex. Then, the recovery strategy is considered as an entire plan or process. In fact, each recovery strategy is composed of many sub-plans or sub-processes. Hence, if we consider the sub-plan factors, the recovery scheme would be more reasonable and detailed.

References

1. Zhao, L.D., Jiang, Y.P.: A replenishment model for emergency rescue systems. In: Proc. of The Third International Conference on Innovative Computing, Information and Control (2008)
2. McDaniels, T., Chang, S., Cole, D., Mikawoz, J., Longstaff, H.: Fostering resilience to extreme events within infrastructure systems: Characterizing decision contexts for mitigation and adaptation. Global Environ. Chang. 18(2), 310–318 (2008)
3. Bruneau, M., Chang, S.E., Eguchi, R.T., et al.: A framework to quantitatively assess and enhance the seismic resilience of communities. Earthq. Spectra. 19(4), 733–752 (2003)
4. Berkes, F., Colding, J.F., Folke, C.: Navigating social-ecological systems: building resilience for complexity and change. Cambridge University Press, UK (2003)
5. Losada, C., Scaparra, M.P., O'Hanley, J.R.: Optimizing system resilience: a facility protection model with recovery time. Eur. J. Oper. Res. 217(3), 519–530 (2012)
6. Zobel, C.W., Khansa, L.: Characterizing multi-event disaster resilience. Comput. Oper. Res. (2011) (in press)
7. Rose, A., Liao, S.Y.: Modeling regional economic resilience to disasters: a computable general equilibrium model of water service disruptions. J. Regional. Sci. 48(1), 75–112 (2005)
8. Nojima, N., Ishikawa, Y., Okumura, T., Sugito, M.: Empirical estimation of lifeline outage time in seismic disaster. In: Proc. of U.S.—Japan Joint Workshop and Third Grantee Meeting, U.S.—Japan Cooperative Research on Urban Earthquake Disaster Mitigation, Seattle, WA, USA, pp. 516–517 (August 2001)
9. Chang, S.E., Rose, A., Shinozuka, M., Svekla, W., Tierney, K.J.: Modeling earthquake impact on urban lifeline systems: advances and integration. In: Research Progress and Accomplishments 1999-2000, MCEER (2000)
10. Bryson, K.M., Millar, H., Joseph, A., Mobolurin, A.: Using formal MS/OR modeling to support disaster recovery planning. Eur. J. Oper. Res. 141(3), 679–688 (2002)
11. Altay, N., Green III, W.G.: OR/MS research in disaster operations management. Eur. J. Oper. Res. 175(1), 475–493 (2006)
12. Sheffi, Y.: Building a resilient supply chain. Harvard Business Review 1(8), 1–4 (2005)
13. Yu, P.L.: A class of solutions for group decision problems. Manage. Sci. 19(8), 936–946 (1973)
14. Zelelny, M.: Compromise programming. In: Cochrane, J.L., Zeleny, M. (eds.) Multiple-Criteria Decision Making, pp. 262–301. University of South Carolina Press, Columbia (1973)
15. Zelelny, M.: A concept of compromise solutions and the method of the displaced ideal. Comput. Oper. Res. 1, 479–496 (1974)
16. Freimer, M., Yu, P.L.: Some new results on compromise solutions for group decision problems. Manage. Sci. 22(6), 688–693 (1976)
17. Goldberg, D.E.: Genetic algorithms in search optimization and machine learning. Addison Wesley, MA (1989)
18. Vose, M.D.: Modelling simple genetic algorithm. In: Foundations of Genetic Algorithms II, pp. 63–73. Morgan Kaufmann Publishers (1993)

A Decision Table Method for Randomness Measurement

Nawaf Alkharboush and Yuefeng Li

Abstract. Data quality has become a major concern for organisations. The rapid growth in the size and technology of a databases and data warehouses has brought significant advantages in accessing, storing, and retrieving information. At the same time, great challenges arise with rapid data throughput and heterogeneous accesses in terms of maintaining high data quality. Yet, despite the importance of data quality, literature has usually condensed data quality into detecting and correcting poor data such as outliers, incomplete or inaccurate values. As a result, organisations are unable to efficiently and effectively assess data quality. Having an accurate and proper data quality assessment method will enable users to benchmark their systems and monitor their improvement. This paper introduces a granules mining for measuring the random degree of error data which will enable decision makers to conduct accurate quality assessment and allocate the most severe data, thereby providing an accurate estimation of human and financial resources for conducting quality improvement tasks.

1 Introduction

Quality assessment is the first and most important step for various applications, including database, data warehouse, data mining, statistics, business management, and knowledge engineering. It plays a major role in the success of these applications. However, quality assessment is the most difficult task to achieve because of massive and continuous data throughput. Therefore, it is considered to be a time consuming, cost excessive and tedious process [19, 23]. Several studies draw the attention to data quality and investigate the impact of poor data on customer satisfaction, decision making, operational costs, and executing strategy [2, 19, 23, 24, 22, 8]. A well

Nawaf Alkharboush · Yuefeng Li
Computer Science Discipline, Faculty of Science and Technology
Queensland University of Technology, QLD 4000, Australia
e-mail: n.alkaharbosh@student.qut.edu.au, y2.li@qut.edu.au

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 13–23.
springerlink.com

known study estimates that immediate cost stemming from the 1-5 percent of error rate is approximately 10 percent of revenue [20].

Recent studies on data quality research have seen significant improvement in the efficient identification and correcting of error data, e.g., incomplete, inaccurate and outliers data. However, the major challenge of data quality research is not just confined by the present and correct of noise data. It includes other major series of steps such as defining, measuring, analysing, and improving data quality. Despite the importance of these four components, most data quality literature circulates their solutions around the first and last steps, defining and improving respectively, of quality problems. Limited attention has been given to data quality assessment, which includes measuring and analysing steps. Contributing to data quality assessment by measuring and analysing poor data can significantly improve data quality research, as it enables decision makers to see the location of the most severe data and determine its patterns. Also, it helps to estimate the time and complexity required to conduct quality improvement tasks.

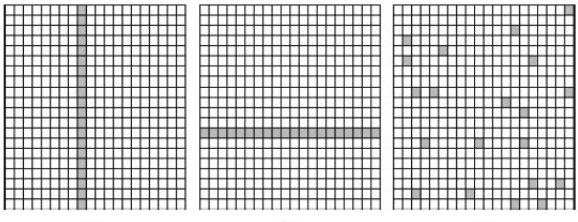


Fig. 1 Error Distributions

There are two elements must be thoroughly investigated to propose a comprehensive data quality assessment method. First, the most common quality assessment measurement relies on error rate or accuracy rate. This often causes interpenetration problems. Error rate divides the total number of defective data to the total fields. Accuracy rate can be calculated by 1-error rate. For example, Figure 1 shows three databases (a), (b) and (c) with 5% of error rate for each one. However, the errors in (a) and (b) are systematically distributed in a column and row respectively, but randomly distributed across columns and rows in database (c) [6]. Undoubtedly, dealing with such errors in databases (a) and (b) is going to be completely different from (c) with regard time, cost, effort, complexity, and solutions.

Secondly, defining the randomness degree of error data for database (c) is an essential step for quality assessment. It enables decision makers to have more insight into quality problems in the database and provides further accuracy estimation of the complexity of the cleaning process. Existing randomness methods presented in [7, 9, 6] investigate the random problems and measure their degree. Yet, these methods encounter critical problems. The first issues is that the randomness methods have a global view of random problems by only measuring the randomness degree

of a table. Yet, they do not consider the fact that a random issue is also independent with its attributes and fields. Another problem is that these methods [7, 9, 6] are complicated and have high time complexity. Finally, users cannot precisely find the potential for most severe quality problems and measure their impact on the quality of a database.

In this paper, we intend to introduce a new approach for assessing data quality by using granules mining. Primarily, this paper will contribute to the following:

1. We formulate the randomness degree problem and introduce a novel measurement for measuring the randomness degree of error data based on granules mining.
2. We propose a decision table method which goes beyond measuring the randomness to distinguish between systematic and random errors.
3. This study considers the fact that there are usually large random patterns with the same minimum supports. Hence we introduce granules taxonomy structure for extracting systematic patterns from random patterns.

The remainder of the paper is structured as follows. We begin by introducing the relevant background in Section 2. In Section 3, we introduce the concept of decision tables for granule mining and randomness measurement. We also discuss the relationship between granule and error distribution. In Section 4, we describe empirical experiments. The last Section includes the conclusion and future work.

2 Related Work

2.1 Current Quality Assessment

For the last decade, researchers and practitioners provided significant contributions to quality assessment. This resulted in the introduction of various methodologies and measurements [19, 18, 4, 3]. Some studies attempted to address the problems of defining a composition algebra for data quality dimensions as in [11, 14, 12, 21]. For example, Paper [11] proposed a model to approximate the similarity between the actual stored data called r-actual and r-ideal for real world database. Another study went further by classifying attributes to identifier attributes and non identifier attributes [14]. In [14], the authors, distinguish between pairs tuples that differ on the identifier attributes, and tuples that are identical in identifier attributes but different in non identifier attributes. A different study investigated the quality composition from the integration system [12, 13]. These methods commonly used error rate or accuracy rate to represent quality probability for dimensions of completeness, accuracy and consistency. However, the problem with existing quality measurement (error rate and accuracy rate) is that users cannot benchmark the quality of its systems. Error rate or accuracy rate might report the same probability but further analysis can show the location of these errors are different.

The lack of regularity or randomness problems has also been investigated in mathematics and compute science. The most well known algorithms for randomness

measurement are called Kolmogorov and Lempel-Ziv (LZ) complexity presented in [7, 9] respectively. Recent study [6] attempted to utilise randomness measurement presented in [9] to measure the randomness degree of errors in a database. The authors [6] follow the common way in quality assessment by transferring the database values to a matrix table [19, 18, 5]. In a matrix table, all values are in (0, 1), where 0 indicates accurate values and 1 indicates poor data regardless of whether its outliers, missing or inaccurate. Then, paper [6] combines two parameters including error rate and poisson distribution with Lempel-Ziv (LZ) random measurement to assess error data. However, these approaches [7, 9, 6] have time complexity and users need to segment the database to small portions to apply the [6] algorithm. Secondly , these randomness methods can only have a global view of the table, but sometime the errors that appear may have different probability distribution.

Other interesting areas for data quality assessment including finding error patterns, determining the systematic and random errors, and allocating the location of the most severe data require more investigation. This paper introduces a new method for data quality assessment based on granule mining. particularly, we contribute to random measurement, distinguish between systematic and random errant patterns, and extract and expose useful error patterns from large random patterns. This will enable decision makers to find the most severe patterns and estimate the right resources needed for solving error patterns.

3 Proposed Method

3.1 Decision Table

Rough set theory formulates the foundation of our quality assessment method. The concept of rough set was proposed by [15]. It has been widely applied in many areas include knowledge discovery, decision analysis, data mining, machine learning, and artificial intelligence. We refer the reader to [15] for fundamental concept of rough set theory and we also suggest [16] which provides a survey analysis for all well know literature that adopt rough set theory with different fields.

Based on rough set theory, users can describe the knowledge in information tables [17]or multi-tier structures [10]. Additionally, users can represent the association among data. Despite the popularity of rough set theory, little research has been conducted in the area of data quality specially quality assessment.

Definition 1. ***(Decision Table)*** *Formally, we call the tuple* (T, A, C, D) *a decision table of database table* T *if* $C \cap D = \emptyset$ *and* $C \cup D \subseteq A$*, where is the set of fields of* T*, or the set of attributes.*

We usually assume that there is a function for every attribute $a \in A$ such that $a : T \rightarrow V_a$, where V_a is the set of all values of a. We call V_a the domain of a, for example, $V_a = \{1, 0\}$ in our examples.

Let B be a subset of A. B determines a binary relation $I(B)$ on T such that $(t_1, t_2) \in I(B)$ if and only if $a(t_1) = a(t_2)$ for every $a \in B$, where $a(t)$ denotes the value of attribute a for object $t \in T$. It is easy to prove that $I(B)$ is an equivalence relation,

and the family of all equivalence classes of $I(B)$, that is a partition determined by B, is denoted by T/B.

The equivalence classes in $T/(C \cup D)$ are also referred to granules; and the equivalence classes in T/C (or T/D) are referred to *C-granules* (or *D-granules*). The equivalence class in $T/(C \cup D)$ (or in T/C, T/D) that contains transaction t is called *granule* (or *C-granule*, *D-granule*) induced by t, and is denoted by $(C \cup D)(t)$ (or $C(t)$, $D(t)$).

This study assumes that there is a database T. Formally, T can be described as a decision table DT_A with (G_i,A_i), where G_i is a set of granules about attributes A_i, and a granule is a group of objects (rows) which has the same attributes' values [15].

Attributes A_i can be divided into two groups, condition attributes C_i and decision attributes D_i, such that $C_i \cap D_i = \emptyset$ and $C_i \cup D_i \subseteq A_i$. Every granule in the decision table can be mapped into a decision rule:

Users can assign condition attributes and decision attributes according to different data quality requirements. For example less error attributes can be the condition and more error attributes can be the decision.

The support of rule $(cg_i \rightarrow dg_i)$ is the $Sup(cg_i \wedge dg_i)$, where $(cg_i \wedge dg_i)$is a granule of Table; and the confidence is

$$Confidence = \frac{Sup(cg_i \wedge dg_j)}{Sup(cg_i)} \tag{1}$$

3.2 New Definition for Randomness Degree

In data quality research, errors are either systematically or randomly distributed across rows and columns. A systematic error presents a clear pattern of defective data. For example, errors might frequently occur on specific columns like address and zip code. On the other hand, randomness distribution of errors show lack of regularity of defective patterns. Intuitively, detecting and handling errors in systematic types are less complicated than randomness errors [6]. Unlike systematic errors, randomness errors are difficult and requires massive amounts of time and cost. The first reason is that finding a value (or values) that produces random patterns is difficult and time consuming. Secondly, the size of errant random patterns is usually large with low frequency. This makes finding potential errant patterns for solving large quality problems very difficult for decision makers to achieve.

Firstly, we introduce the randomness measurement based on patterns or (granules). Then, we further enhance the proposed solution by measuring the distribution of granules by two vectors: numbers of patterns or (granules) and the patterns distribution. This enables users to expose and assign systematic and random patterns to different categories.

In this section we measure the randomness degree of errors based on the numbers of errant patterns or errant granules. For example, Table 1 and Table 2 are two decision tables DT_{A1} and DT_{A2} respectively generated from two different tables that have the same attributes "a_1,a_2,a_3" and the same size 100 rows, but with different errors distributions.

Table 1 Cover All Patterns

Granule	a_1	a_2	a_3	$Sup(g_i)$
g_1	0	0	0	50
g_2	0	0	1	25
g_3	0	1	1	2
g_4	0	1	0	2
g_5	1	1	1	2
g_6	1	0	0	2
g_7	1	1	0	15
g_8	1	0	1	2

Table 2 Not Cover All Patterns

Granule	a_1	a_2	a_3	$Sup(g_i)$
g_1	0	0	0	50
g_2	0	0	1	25
g_3	0	1	1	10
g_4	0	1	0	15

As can be seen from DT_{A1} in Table 1 and DT_{A2} in Table 2, DT_{A1} has more errant granules or patterns than DT_{A2}. This indicates that DT_{A1} presents more random errors than DT_{A2}. Based on the definition of randomness degree presented in paper [1], we can calculate the randomness degree RD_A of errant data as follows:

Definition 2. ***(Randomness Degree)*** *Let $|DT_A|$ is the number of errant granules g_i in DT_A and $2^{|Ai|}-1$ is the size of covering set for errant granules in DT_A. The random degree is defined as:*

$$RD_A = \frac{|DT_A|}{2^{|Ai|}-1} \tag{2}$$

The size of the covering set includes normal patterns (all A_i fields are 0) and numbers of errant patterns. For example, the size of the covering set for DT_{A1} in Table 1 is $2^{|3|} = 8$. This covers a normal pattern and all possible errant patterns generated from a database regardless of its number or rows. Similarly, the size of covering set for DT_{A2} Table 2 is the same as Table 1 because both tables have the same attributes $|A_i| = 3$. Hence, users need to exclude a normal pattern when they calculate randomness degree RD_A.

Referring back to the DT_{A1} in Table 1 and DT_{A2} in Table 2, we can apply Eq.(2) to measure their RD_A. For example, DT_{A1} in Table 1 shows seven errant patterns, e.g., g_2 to g_8. By using the Eq.(2), the RD_A of error in DT_{A1} Table 1 is $7/7 = 100\%$. This means that errors values cover all possible errant patterns in T. However, in some scenarios, the errant patterns g_i in a decision table is less than the covering set as in DT_{A2} Table 2. Therefore, we have low RD_A $3/7 \approx 43\%$.

3.3 Distinguish between Systematic and Random Patterns

In this part, we attempt to discriminate between the systematic and random patterns. Previously, we utilise Eq.(2) to calculate the randomness degree; and if RD_A $>$ threshold, then errors in T have random distribution. For example, if we assume the threshold for the above decision tables Table 1 and Table 2 is 50%, then the DT_{A1} in Table 1 has random errors with 100% random degree and Table 2 has

Table 3 Distinguish between Systematic and Randomness Errors

Error Rate	RD_A	*Number of*	*Total (Sup)*	*Systematic Errant Patterns*		*Improve*	*Random Errant Patterns*		*Improve*	*Patterns*
(%)	*(%)*	(g_i)	*of* (g_i)	*N. Patterns*	*Support*	*quality (%)*	*N. Patterns*	*Support*	*quality (%)*	*Impact quality*
2	*2.1*	*216*	*3132*	18	2618	83.6	0	0	0	S
5	*6.2*	*644*	*6159*	103	4998	81.1	40	390	6.3	S
10	*17.7*	*8698*	*8698*	196	6103	71.1	105	445	5.1	S
20	*51.5*	*5296*	*10096*	734	4835	47.9	4562	5261	52.1	S/R
35	*92.5*	*9505*	*10275*	84	279	2.7	9421	9996	97.3	R
50	*98.2*	*10094*	*10281*	3	9	0.08	10091	10272	99.91	R
65	*92.2*	*9480*	*10281*	89	292	2.9	9391	9989	97.2	R
80	*51.5*	*5294*	*10281*	747	5018	48.8	4547	5263	51.2	R/S
90	*17.8*	*1831*	*10281*	178	7577	73.7	41	214	2.1	S

systematic errors. However, when we analyse DT_{A1} 1, we notice that some errors occur more frequently in some patterns. This leads us to a critical question, is measuring randomness degree enough for a complete quality assessment solution?

To obtain deep insight into this problem, we examine three different decision tables with different distribution of errors; see Figure 2. Be reminded that we only change the support on defective patterns to have different distribution weights, while the RD_A degree remains the same at 100%. Although, all three distributions A, B and C have the same RD_A with 100% randomness degree, handling some errors in distribution like A is less complicated than others as in distribution C. Also, fixing several patterns e.g. (g_2 and g_7) in A could significantly improve the quality of a database by a large percentage. Hence, we enhance the RD_A presented in this paper by introducing two measurements to categorise systematic patterns and random patterns into two groups. Also, we calculate the impact of each pattern on the quality. This enables decision makers to target the most serious errant patterns based on their time and resources.

The first measurement is based on the pattern support. In a decision table, each pattern has a support which indicates the frequency of this type of error occurring in a database. In this way, users can specify a threshold to find the severity of defective patterns. If $Sup(gi) >$ threshold, the nominated pattern is considered to be a systematic pattern. For example, if users specify 8 as the minimum threshold for

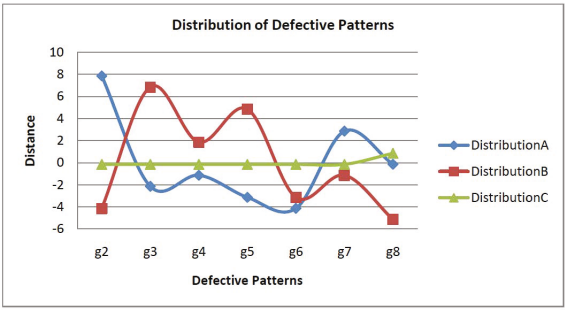


Fig. 2 Distance Distribution

Distribution A, then we have two severe errant patterns, g_2 and g_7. Also, this solution can be applied to distribution B and C but users need to modify the minimum threshold accordingly.

The second distinguishing measurement between systematic and random patterns is based on the distance between patterns. In this type of measurement, we measure the distance between defective patters to allocate the location of systematic S and random patterns R. More importantly, we distiguesh between pure random patterns PR and weak random patterns WR.

Definition 3. ***(Define Systematic and Random Patterns)*** *Gevin a granule distance dis_{g_i} and the minimum distance min_dis_1, the systematic granule (S) and Random granule (R) are defined as:*

$$g_i = \begin{cases} S \;\; if \;\; dis_{g_i} > min\_dis_1 \\ R \quad = \begin{cases} PR \;\; if \;\; dis_{g_i} > min\_dis_2 \\ WR \;\; if \;\; dis_{g_i} < min\_dis_2 \end{cases} \end{cases}$$

To measure the distance, we firstly need to calculate the average support avg_sup of defective patterns. Then, we calculate the distance value dis_{g_i} between a patterns $Sup(gi)$ and avg_sup.

$$avg\_sup = \frac{1}{|DT_A|} \sum_{g \in DT_A} Sup(g_i) \tag{3}$$

We determine the severity of systematic or randomness errors based on the numbers of defective patterns on each category and their supports. For example, Figure 2 depicts the distance distributions for A, B and C. From this graph, we can easily determine which pattern is far from the average support avg_sup and which is not. However, Distribution C has the liner distribution of errant patterns. This indicates that these errant patterns mostly have random errors.

In this study, we call (systematic S and pure random patterns PR) as useful patterns (U) for solving potential quality problems. Also, we consider (weak random errant patterns WR) as outlier patterns and labelled as less useful patterns for quality problems. However, we highlight and measure the impact of both categories, useful (U) and less useful errant patterns (L), on quality improvement. Eq.(4) calculates the quality improvement that user could obtain from correcting only the useful patterns.

$$quality\_improvement = \frac{\sum_{g \in U} Sup(g)}{\sum_{g \in DT_A} Sup(g)} \tag{4}$$

where $\sum_{g \in U} Sup(g)$ is the total support of useful patterns and $\sum_{g \in DT_A} Sup(g)$ is the total support of all defective patterns. This equation can also be used to calculate the the quality improvement for less useful patterns.

4 Empirical Experiments

4.1 Empirical Design

The decision table method for quality assessment has been evaluated based on a real database (FoodMart 2005) available from `http://www.microsoft.com`. This database consists of a number of tables. We examine the efficiency and effectiveness of our approach on Customer relation. In the customer table, there are 10281 customers (rows) and 18 attributes such as Customer-id, Name, Gender, Zip Code, Income,..., and Education. We transfer the actual values to the matrix table by assigning 0 value to normal data and 1 value for abnormal including (missing, outliers, incomplete)data because our scope in this paper excludes finding missing or outliers. This study rather addresses a critical problem for quality assessment literature like Randomness measurement.

4.2 Results and Discussions

We randomly and gradually increase the error rate ten times be various percentage as in Figure 4. Then, the decision table was constructed for each table. This means we have ten different decision tables with different random degree of errors as in Figure 4.

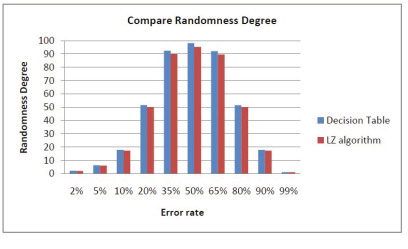


Fig. 3 Compare Randomness Degree

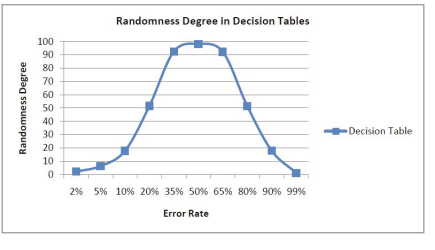


Fig. 4 Randomness Degree in DT

As can be seen in Figure 4, the numbers of defective patterns increase gradually among 2%,5%,10%,20% and 35%. The increase of error rate by 2%,5%,10%,20% and 35% contributes to the increase of the randomness degree from 2.1% to up 92.4%. Reversely, the randomness degree decreases when the distribution of errors is greater than 50%. The randomness degree reaches the maximum level when the error rate is 50%.

To certify the validity of the experiment results, we rigorously compare the results obtained from our algorithm with the LZ algorithm [9, 6]. We compare decision table algorithm and LZ algorithm along different error distribution rates 2%,5%,10%,20%,35%,50%,65%,80%,90%,99%. The results shown in Figure 3 conclusively demonstrate that the decision table method provides similar results

to the LZ algorithm but with new method. Additionally, our algorithms is fast and does not have time complexity like the LZ algorithm.

Another signification contribution of our approach is that users can distinguish between defective patterns that tend to have a systematic attitude from the random one, see Table 3. After we measure the randomness degree RD_A, we classify defective errant patterns into two groups, systematic and random. Then, we calculate the impact of systematic and random errant on the quality by dividing the total support of systematic or random patterns to the total support of all errant patterns. This enables users to determine which type of errors (systematic (S) or random (R))have the most impact on the quality and thereby determine the potential patterns for improving data quality, Table 3.

5 Conclusion and Future Work

Data quality assessment is a critical component for organisational performance as it supports decision makers meeting organisational needs. This paper has introduced a novel approach based on granule mining for data quality assessment. Primarily, we have contributed to solving three major issues in data quality assessment. Firstly, we present a new and efficient algorithm for measuring the randomness degree of error data. Unlike other randomness measurements [7, 9, 6], our algorithm is fast and does not have time complexity problems. We also solved the problem of determining between systematic and random errant patterns and revealing their impact on the quality. The empirical studies presented in this paper have shown promising results for granules mining in measuring randomness degree, defining systematic and random errors and reducing the size of errant random patterns.

There are several interesting directions we are considering in future. (1) Using granule mining for finding and correcting various types of data quality dimensions, and (2) Utilising granules for quality improvement.

References

1. Alkharboush, N., Li, Y.: A decision rule method for data quality assessment. In: Proceedings of the 15th International Conference on Information, vol. 3, pp. 84–95. ACM, Little Rock (2010)
2. Ballou, D.P., Pazer, H.L.: Modeling data and process quality in multi-input, multi-output information systems. Management Science 31(2), 150–162 (1985)
3. Batini, C., Scannapieco, M.: Data quality: Concepts, methodologies and techniques. Springer-Verlag New York Inc. (2006)
4. Dasu, T., Johnson, T.: Exploratory data mining and data cleaning. Wiley, New York (2003)
5. Even, A., Shankaranarayanan, G.: Dual assessment of data quality in customer databases. Journal of Data and Information Quality 1(3), 1–29 (2009)
6. Fisher, W.C., Lauria, J.M.E., Matheus, C.C.: An accuracy metric: Percentages, randomness, and probabilities. Journal of Data and Information Quality 1(3), 1–21 (2009)

7. Kolmogorov, A.: Three approaches to the quantitative definition of information. Problems of Information Transmission 1(1), 1–7 (1965)
8. Lee, Y.W., Strong, D.M., Kahn, B., Wang, R.Y.: Aimq: a methodology for information quality assessment. Information & Management 40(2), 133–146 (2002)
9. Lempel, A., Ziv, J.: On the complexity of finite sequences. IEEE Transactions on Information Theory 22(1), 75–81 (1976)
10. Li, Y.: Interpretations of discovered knowledge in multidimensional databases. In: Proceedings in IEEE International Conference on Granular Computing, p. 307 (2007), doi:10.1109/GrC.2007.92
11. Motro, A., Rakov, I.: Estimating the quality of databases. Flexible Query Answering Systems, 298–307 (1998)
12. Naumann, F., Freytag, J., Leser, U.: Completeness of integrated information sources. Information Systems 29(7), 583–615 (2004)
13. Parssian, A.: Managerial decision support with knowledge of accuracy and completeness of the relational aggregate functions. Decision Support Systems 42(3), 1494–1502 (2006)
14. Parssian, A., Sarkar, S., Jacob, V.: Assessing data quality for information products: Impact of selection, projection, and cartesian product. Management Science 50(7), 967–982 (2004)
15. Pawlak, Z.: Rough sets: Theoretical aspects of reasoning about data. Springer, Kluwer, Dordrecht (1991)
16. Pawlak, Z., Skowron, A.: Rough sets and boolean reasoning. Information Sciences 177(1), 41–73 (2007)
17. Pawlak, Z., Skowron, A.: Rudiments of rough sets. Information Sciences 177(1), 3–27 (2007)
18. Pipino, L.L., Lee, Y.W., Wang, R.Y.: Data quality assessment. Communications of the ACM 45(4), 211–218 (2002)
19. Redman, T.C.: Data quality for the information age. Artech House Boston, MA (1996)
20. Redman, T.C.: The impact of poor data quality on the typical enterprise. Communications of the ACM 41(2), 79–82 (1998)
21. Scannapieco, M., Batini, C.: Completeness in the relational model: a comprehensive framework. In: Proceedings of 9th International Conference on Information Quality, ICIQ, vol. 4, pp. 333–345 (2004)
22. Strong, D.M., Lee, Y.W., Wang, R.: Data quality in context. Communications of the ACM 40(5), 103–110 (1997)
23. Wand, Y., Wang, R.: Anchoring data quality dimensions in ontological foundations. Communications of the ACM 39(11), 86–95 (1996)
24. Wang, R.W., Strong, D.: Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems 12(4), 5–33 (1996)

A Descriptor-Based Division Chart Table in Rough Non-deterministic Information Analysis

Hiroshi Sakai, Hitomi Okuma, Mao Wu, and Michinori Nakata

Abstract. We have been proposing a framework *Rough Non-deterministic Information Analysis* (*RNIA*), which handles rough set-based concepts in not only *Deterministic Information Systems* (*DISs*) but also *Non-deterministic Information Systems* (*NISs*). We have recently suggested a new visual notation named *Division Chart* over an equivalence class defined by descriptors. In this paper, we extend it further towards a *Division Chart Table* (*DCT*). *DCTs* are useful for modeling rough set and granular computing concepts, in particular consistency-based foundations of rule generation and data dependency. They also provide more comprehensive interpretation of the contents of *NISs* than methods proposed with this respect in the previous research related to *RNIA*.

1 Introduction

Rough set theory offers a mathematical approach to vagueness and uncertainty, and the rough set-based concepts have been recognized to be very useful [10, 2, 6, 19].

Hiroshi Sakai
Department of Basic Sciences, Faculty of Engineering, Kyushu Institute of Technology, Tobata, Kitakyushu 804-8550, Japan
e-mail: sakai@mns.kyutech.ac.jp

Hitomi Okuma
Faculty of Education and Welfare Science, Oita University Dannoharu, Oita 870, Japan
e-mail: okuma@oita-u.ac.jp

Mao Wu
Department of Integrated System Engineering, School of Engineering, Kyushu Institute of Technology, Tobata, Kitakyushu 804-8550, Japan

Michinori Nakata
Faculty of Management and Information Science, Josai International University Gumyo, Togane, Chiba 283, Japan
e-mail: nakatam@ieee.org

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 25–34.
springerlink.com

This theory usually handles tables with deterministic information, which we call *Deterministic Information Systems* (*DISs*). Many applications of this theory to information analysis, data mining, rule generation, machine learning and knowledge discovery have been investigated [10, 2, 6, 11, 12, 18, 19].

Non-deterministic Information Systems (*NISs*) and *Incomplete Information Systems* (*IISs*) have been proposed for handling information incompleteness in *DISs* [1, 3, 4, 5, 8, 9]. *NISs* have been recognized to be the most important framework for handling information incompleteness in tables, and several theoretical works have been reported. We follow this robust framework, and we have been developing algorithms and a software tool, which can handle rough set-based concepts in *NISs*. We are calling this work *Rough Non-deterministic Information Analysis* (*RNIA*).

Recently, we proposed a notation named *Division Chart* [7], and clarified the relation between inclusion relations on classes and the consistency. In this paper, we briefly refer to *RNIA* and division charts, then propose *Division Chart Tables* for *RNIA*. This division chart table gives us comprehensive interpretation to the previous research about *RNIA*.

The paper is organized as follows: Section 2 recalls a mathematical framework *RNIA* and issues by an example, and Section 3 proposes division charts and a division chart table including the enhancement of the proposition on the consistency. Section 4 recalls *inf* and *sup* blocks instead of an equivalence class. Then, we propose division charts and a division chart table in *NISs*, and clarify their properties. Section 5 concludes the paper.

2 Issues on Rough Non-deterministic Information Analysis

Let us consider Table 1 and Table 2. Table 1 is an exemplary *DIS* ψ_1, and Table 2 is an exemplary *NIS* Φ_1.

Table 1 A Deterministic Information System ψ_1.

patient	*headache*	*temperature*	*flu*
1	*yes*	*very_high*	*yes*
2	*yes*	*very_high*	*yes*
3	*no*	*normal*	*no*

Table 2 A Non-deterministic Information System Φ_1.

patient	*headache*	*temperature*	*flu*
1	{*yes*}	{*very_high*}	{*yes*}
2	{*yes*, *no*}	{*high*, *very_high*}	{*yes*}
3	{*no*}	{*normal*, *high*}	{*yes*, *no*}

In a *DIS*, each attribute value is fixed. However, in a *NIS* each attribute value is given as a set. We interpret this set as that the actual value exists in this set but it is unknown due to the information incompleteness. *NISs* were proposed by Pawlak, Orłowska and Lipski in order to handle information incompleteness in *DISs* [4, 5, 8, 9].

In *NISs*, the concept of an *extension* has been employed. Namely, it is possible to generate a *DIS* by replacing every set of values with a value in the set. We name such *DISs derived DISs* from a *NIS*. In Φ_1, there are 16 derived *DISs*, and ψ_1 is a derived *DIS* from Φ_1. The following two modalities *certainty* and *possibility* are introduced into *NISs*.

(Certainty) If a formula F holds in each derived *DIS* from a *NIS*, F also holds in the unknown real *DIS*.

(Possibility) If a formula F holds in some derived *DISs* from a *NIS*, there exist such a possibility that F holds in the unknown real *DIS*.

We have coped with several issues related to these two modalities [14, 15, 16, 17]. The most important problem is how to compute two modalities depending upon all derived *DISs* from a *NIS*. The number of all derived *DISs* increases in exponential order, therefore a simple method, such that every definition is sequentially computed in all derived *DISs* from a *NIS*, is not suitable. We have proposed some rough set-based methods [15, 16] for solving this problem.

3 Division Charts and Division Chart Tables in DISs

3.1 Foundations of DISs

A *Deterministic Information System* (*DIS*) is a quadruplet $(OB, AT, \{VAL_A | A \in AT\}, f)$, where OB is a finite set whose elements are called *objects*, AT is a finite set whose elements are called *attributes*, VAL_A is a finite set whose elements are called *attribute values* and f is such a mapping that $f : OB \times AT \rightarrow \cup_{A \in AT} VAL_A$.

We usually suppose condition attributes CON and decision attributes DEC. An object x is *consistent* for CON and DEC, if $f(x,CON)=f(y,CON)$ $(y \neq x)$ means $f(x,DEC)=f(y,DEC)$.

In a *DIS*, we usually employ a pair $[A, val_A]$ $(A \in AT, val_A \in VAL_A)$, and call this pair a *descriptor*. Let $[x]_A$ denote an equivalence class $\{y \in OB | f(y,A)=f(x,A)\}$. For CON, $[x]_{CON}$ is defined by $\cap_{A \in CON}[x]_A$.

3.2 A Division Chart in DISs

For an object x satisfying $f(x,CON)=p$ and $f(x,DEC)=q$, we consider a division over an equivalence class $[x]_{[CON,p]}$ by a descriptor $[DEC,q]$. We employ $[x]_{[CON,p]}$ instead of $[x]_{CON}$ in order to clarify the attribute value p from now on. In this case, we can obtain either τ_1 or τ_2 from each object in $[x]_{[CON,p]}$.

$\tau_1 : [CON,p] \Rightarrow [DEC,q], \qquad \tau_2 : [CON,p] \Rightarrow [DEC,q'](q' \neq q).$

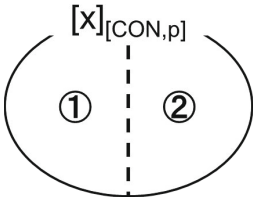


Fig. 1 A division chart $DC(p,q)$ over $[x]_{[CON,p]}$ by $[DEC,q]$ in a *DIS*.

By using definable implications, we can give a division over $[x]_{[CON,p]}$ in Fig. 1. Here,

①=$\{y \in [x]_{[CON,p]} \mid f(y,CON) = p,\ f(y,DEC) = q\}$,
②=$\{y \in [x]_{[CON,p]} \mid f(y,CON) = p,\ f(y,DEC) = q'\ (q' \neq q)\}$,

We name a diagram expressing such division a *division chart* $DC(p,q)$ of $[x]_{[CON,p]}$ by $[DEC,q]$. We call ① and ② *components* of $[x]_{[CON,p]}$. Due to $DC(p,q)$, we enhance the proposition on consistency [10] to the following.

Proposition 1. For every *DIS*, (1), (2) and (3) in the following are equivalent.
(1) Each object $y \in [x]_{[CON,p]}$ $(\neq \emptyset)$ is consistent for $[CON,p]$ and $[DEC,q]$.
(2) $[x]_{[CON,p]} \subseteq [x]_{[DEC,q]}$.
(3) The component ② in $DC(p,q)$ is an empty set.

(Proof.) (1) and (2) are equivalent due to [10].
((2)⇒(3)) Since any $y \in [x]_{[CON,p]}$ satisfies $y \in [x]_{[DEC,q]}$, $f(y,DEC)$=q. Therefore, any y defines an implication $[CON,p] \Rightarrow [DEC,q]$. This means ②=$\emptyset$.
((3)⇒(2)) Let us suppose $y \in [x]_{[CON,p]}$ and $y \notin [x]_{[DEC,q]}$. Then, $f(x,CON)$=p and $f(y,DEC)$=q' $(q \neq q')$, and y becomes an element of ②. This means ② $\neq \emptyset$, and contradicts the assumption.

3.3 A Division Chart Table in DISs

Since $DC(p,q)$ depends on each pair of $([CON,p],[DEC,q])$, we can consider a set of all $DC(p,q)$. We name this set of all $DC(p,q)$ a *division chart table* of a *DIS* ψ. Let $DCT(\psi,CON,DEC)$ denote it, and we may employ the notation $①_{p,q}$ for the component ① in $DC(p,q)$. We have simply implemented this software tool, and the following is the logging data for obtaining $DCT(\psi_1,[1,2],3)$ in Table 1.

```
?-dct([1,2],3). /* CON:[1,2], DEC:3 */
dc(1,[yes,normal,yes],[],[]). /* the 2nd:Implication p⇒q */
:     :     :
dc(5,[yes,very_high,yes],[1,2],[]). /* the 3rd:Component ① */
dc(6,[yes,very_high,no],[],[1,2]). /* the 4th:Component ② */
```

```
dc(7,[no,normal,yes],[],[3]).
dc(8,[no,normal,no],[3],[]).
:     :     :
dc(12,[no,very_high,no],[],[]).
EXEC_TIME=0.0(sec)
yes
```

Here, condition attributes are the 1st (*headache*) and the 2nd (*temperature*). The decision attribute is the 3rd (flu). This is specified by a command $dct([1,2],3)$. From the 5th and the 8th division charts, we know each object is consistent, because ① $\neq \emptyset$ and ②=$\emptyset$ ((3) of Proposition 1) hold. Furthermore, we know the degree of dependency is 1 (=3/3). We have the application of $DCT(\psi, CON, DEC)$ in the following.

Remark 1. Suppose CON and DEC in a DIS ψ.
(1) A task of consistency-based rule generation is to pick up the 2nd argument in a division chart satisfying $①_{p,q} \neq \emptyset$ and $②_{p,q}$=$\emptyset$.
(2) The degree of dependency is $|\cup_{p,q}\{x \in ①_{p,q} | ②_{p,q} = \emptyset\}|/|OB|$.

After obtaining $DCT(\psi, CON, DEC)$, we can calculate the above by one time search over $DCT(\psi, CON, DEC)$. Since the most of the necessary calculation on consistency have been done in this table, $DCT(\psi, CON, DEC)$ means the summary of consistency for CON and DEC.

4 Division Charts and Division Chart Tables in NISs

4.1 Foundations of NISs

A *Non-deterministic Information System* (*NIS*) is a quadruplet $(OB, AT, \{VAL_A | A \in AT\}, g)$ [8, 9]. OB, AT and VAL_A are finite sets like in *DISs*, and g is a mapping from $OB \times AT$ to $P(\cup_{A \in AT} VAL_A)$ (a power set of $\cup_{A \in AT} VAL_A$). Every set $g(x,A)$ is interpreted as that there is an actual value in it but it is not known.

For a $NIS = (OB, AT, \{VAL_A | A \in AT\}, g)$ and a set $ATR \subseteq AT$, we name a $DIS = (OB, ATR, \{VAL_A | A \in ATR\}, h)$ satisfying $h(x,A) \in g(x,A)$ a *derived DIS* (for ATR from a *NIS*). In Φ_1 in Table 2, there are 16 (=2^4) derived *DISs*. For a *NIS* Φ, let $DD(\Phi)$ denote a set of all derived *DISs*.

4.2 Two Blocks inf and sup for Descriptors

We employed each equivalence class $[x]_{[A,val_A]}$ (=$\{y \in OB | f(y,A)$=$f(x,A)$=$val_A\}$) in *DISs*. However, we do not have to specify val_A in a *DIS*, because $f(x,A)$=val_A uniquely holds. The specification of val_A is redundant in a *DIS*.

On the other hand in *NISs*, $[x]_A$ is meaningless, because it depends upon derived *DISs*, and we need to specify $[x]_{[A,val_A]}$ instead of $[x]_A$. Here, we define two *blocks*

inf and sup. The usage of each block by Grzymała-Busse is known well [1], and we enhance the definition of a block.

Definition 1. In a $NIS = (OB, AT, \{VAL_A | A \in AT\}, g)$, we define the following two sets of objects, i.e., inf and *sup blocks*, for each descriptor $[A, val_A]$ ($A \in ATR \subseteq AT$, $val_A \in VAL_A$).
(1) $inf([A, val_A]) = \{x \in OB | \ g(x,A) = \{val_A\}\}$,
(2) $inf(\wedge_{A \in ATR}[A, val_A]) = \cap_{A \in ATR}\ inf([A, val_A])$,
(3) $sup([A, val_A]) = \{x \in OB | \ val_A \in g(x,A)\}$,
(4) $sup(\wedge_{A \in ATR}[A, val_A]) = \cap_{A \in ATR}\ sup([A, val_A])$.

Now, let us consider a relation between a class $[x]_{[A,val_A]}$ and two blocks inf and sup. If each attribute value in a NIS is a singleton set, we can see it a DIS and we derive the following:

$[x]_{[A,val_A]}$=$inf([A, val_A])$=$sup([A, val_A])$ ($f(x,A)$=val_A).

However in every NIS, $[x]_{[A,val_A]}$ depends upon derived $DISs$, and the following holds.

$inf([A, val_A]) \subseteq [x]_{[A,val_A]} \subseteq sup([A, val_A])$.

We employed these inf and sup blocks, and solved several issues [14, 15, 16, 17].

4.3 A Division Chart in NISs

In $NISs$, a block $sup([CON, p])$ is the maximum set for $[x]_{[CON,p]}$, so we consider a division of $sup([CON, p])(\neq \emptyset)$ by a descriptor $[DEC, q]$. In this case, we can characterize obtainable implications in each component in Table 3, and we have the following division chart in Fig. 2. Since $inf([CON, p]) \subseteq sup([CON, p])$ holds, a block $inf([CON, p])$ is also divided.

4.4 A Division Chart under Some Inclusion Relations

Now, this sub-section considers two variations of Proposition 1. At first we need to consider the consistency of objects by descriptors $[CON, p]$ and $[DEC, q]$. For a NIS Φ, an object x and two descriptors $[CON, p]$, $[DEC, q]$ satisfying $x \in sup([CON, p]) \cap sup([DEC, q])$, we define the following:

Table 3 Six components of $sup([CON, p])$. Here, $p \Rightarrow q$ means $[CON, p] \Rightarrow [DEC, q]$.

Component	*CON*	*DEC*	*Obtainable implications*
①	$\{p\}$	$\{q\}$	$p \Rightarrow q$
②	$\{p\}$	*a set containing q*	$p \Rightarrow q, p \Rightarrow q'$ $(q \neq q')$
③	$\{p\}$	*a set not containing q*	$p \Rightarrow q'$
④	*a set containing p*	$\{q\}$	$p \Rightarrow q, p' \Rightarrow q$ $(p \neq p')$
⑤	*a set containing p*	*a set containing q*	$p \Rightarrow q, p \Rightarrow q', p' \Rightarrow q, p' \Rightarrow q'$
⑥	*a set containing p*	*a set not containing q*	$p \Rightarrow q', p' \Rightarrow q'$

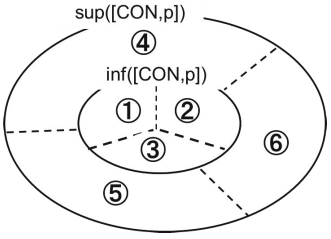


Fig. 2 A division chart of $sup([CON,p])$ by $[DEC,q]$ in a *NIS*. A set containing Y means a set including an element Y and another one element at least.

$DD(\Phi,x,p,q)=\{\psi|\,\psi$ *is a derived DIS satisfying* $h(x,CON)=p,\ h(x,DEC)=q\}$. If $g(x,CON)=\{p\}$ and $g(x,DEC)=\{q\}$, $DD(\Phi,x,p,q)=DD(\Phi)$ holds. However, generally $DD(\Phi,x,p,q)\subseteq DD(\Phi)$ holds. For $CON=\{headache,temperature\}$ and $DEC=\{flu\}$ in Table 2, $|DD(\Phi_1)|=16$ and $|DD(\Phi_1,2,(yes,very\_high),yes)|=4$ hold. We have already characterized the following concepts by inclusion relations on inf and sup [15].

(1) An object x is *globally consistent*, if x is consistent in each $\psi\in DD(\Phi,x,p,q)$.
(2) An object x is *marginal*, if x is consistent in some $\psi\in DD(\Phi,x,p,q)$.
(3) An object x is *globally inconsistent*, if x is inconsistent in each $\psi\in DD(\Phi,x,p,q)$.

Like Proposition 1, we enhance the previous result by division charts.

Proposition 2. For every *NIS*, (1), (2) and (3) in the following are equivalent.
(1) An object $x\in sup([CON,p])\cap sup([DEC,q])$ $(\neq\emptyset)$ is globally consistent for $[CON,p]$ and $[DEC,q]$.
(2) $sup([CON,p])\subseteq inf([DEC,q])$.
(3) Components $②_{p,q}$, $③_{p,q}$, $⑤_{p,q}$ and $⑥_{p,q}$ are all empty sets.

(Proof.) (1) and (2) are equivalent due to [15].
((2)⇒(3)) The proof is in Proposition 2 in [7].
((3)⇒(1)⇔(2)) Due to the assumption, $DC(p,q)$ consists of $①_{p,q}$ and $④_{p,q}$. All obtainable implications are $p\Rightarrow q$ and $p'\Rightarrow q$, therefore there is no contradiction in any case. Thus, an object x is globally consistent.

Proposition 3. For every *NIS*, (1), (2) and (3) in the following are equivalent.
(1) An object $x\in sup([CON,p])\cap sup([DEC,q])$ $(\neq\emptyset)$ is globally inconsistent for $[CON,p]$ and $[DEC,q]$.
(2) $inf([CON,p])\not\subseteq sup([DEC,q])$.
(3) A component $③_{p,q}\neq\emptyset$.

(Proof.) (1) and (2) are equivalent due to [15].
((2)⇒(3)) The proof is in Proposition 6 in [7].
((3)⇒(1)⇔(2)) Due to the assumption, we obtain an implication $p \Rightarrow q'$ in $③_{p,q}$. This implication contradicts $p \Rightarrow q$.

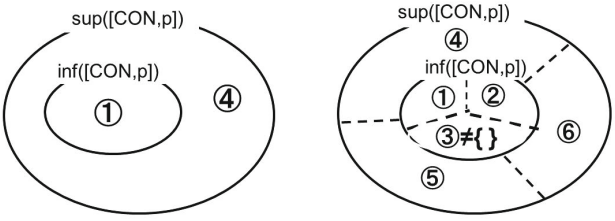


Fig. 3 Revised division charts in Proposition 2 (left side) and Proposition 3 (right side).

4.5 Calculation of Criterion Values by a Division Chart in NISs

In a *DIS*, we usually employ $support(\tau)$ and $accuracy(\tau)$ for evaluating an implication τ. In a *NIS*, these values depend on derived *DISs*, therefore we defined the minimum and the maximum of them. We express them $minsupp(\tau)$, $minacc(\tau)$, $maxsupp(\tau)$ and $maxacc(\tau)$, and have already obtained the calculation of them by inf and sup blocks [16].

It is also possible to consider this calculation by $DC(p,q)$. Generally, $N/M \leq (N+1)/(M+1)$ holds for integers N $(N \geq 0)$, M $(M > 0)$ and $N \leq M$. In order to increase $accuracy(\tau)$, we try to generate an implication $p \Rightarrow q$ from each object, because this object is counted to both denominator and numerator. Thus we can obtain $maxacc(\tau)$. At the same time, we obtain $maxsupp(\tau)$, because the implication $p \Rightarrow q$ occurs maximally. In order to decrease $accuracy(\tau)$, we try to generate an implication $p \Rightarrow q'$ from each object, because this object is counted to only denominator. Like this we can visually know the following formula.

(1) $minsupp(\tau)$=$|①|/|OB|$.
(2) $minacc(\tau)$=$|①|/(|①|+|②|+|③|+|⑤|+|⑥|)$.
(3) $maxsupp(\tau)$=$(|①|+|②|+|④|+|⑤|)/|OB|$.
(4) $maxacc(\tau)$=$(|①|+|②|+|④|+|⑤|)/(|①|+|②|+|③|+|④|+|⑤|)$.

4.6 A Division Chart Table in NISs

The following is the logging data for obtaining $DCT(\Phi_1,[1,2],3)$ in Table 2.

```
?-dct([1,2],3). /* CON:[1,2], DEC:3 */
dc(1,[yes,normal,yes],[],[],[],[],[],[]). /* the 2nd:p⇒q */
:     :     :
```

```
dc(5,[yes,very_high,yes],[1],[],[],[2],[],[]).
dc(6,[yes,very_high,no],[],[],[1],[],[],[2]).
dc(7,[no,normal,yes],[],[],[],[],[3],[]).
dc(8,[no,normal,no],[],[],[],[],[3],[]).
dc(9,[no,high,yes],[],[],[],[2],[3],[]).
:     :     :
dc(12,[no,very_high,no],[],[],[],[],[],[2]]).
EXEC_TIME=0.0(sec)
yes
```

We focus on the 5th $DC((yes, very\_high), yes)$. Since components ②, ③, ⑤ and ⑥ are all empty sets, we apply Proposition 2 and we know the object 1 is always consistent in $\psi \in DD(\Phi_1, 1, (yes, very\_high), yes)$=$DD(\Phi_1)$. Namely, object 1 related to $(yes, very\_high, yes)$ is definite ($g(1, headache)$, $g(1, temperature)$ and $g(1, flu)$ are all singleton), and consistent in all 16 derived *DISs*. The object 2 is also consistent in each $\psi \in DD(\Phi_1, 2, (yes, very\_high), yes) \subset DD(\Phi_1)$. This object 2 related to $(yes, very\_high, yes)$ is not definite, and consistent in each of 4 derived *DISs*. Therefore, object 1 is definite globally consistent, and object 2 is indefinite globally consistent.

This directly connected with rule generation, namely, an implication in the 2nd argument $[yes, very\_high, yes]$, i.e.,

$$[headache, yes] \wedge [temperature, very\_high] \Rightarrow [flu, yes]$$

in the 5th division chart means a rule which is always consistent. Like this we will be able to obtain lots of useful information from $DCT(\Phi, CON, DEC)$.

5 Concluding Remarks

We proposed division charts over a *DIS* and a *NIS*, and extended them to a division chart table in a *DIS* and a *NIS*. We have solved several issues by using inf and sup blocks, but division charts and a division chart table will also take the role of inf and sup blocks. Since a division chart table is visual and comprehensive, we are revising previous complicated proofs by it. The execution logs of experiments including $DCT s$ by the implemented software tools are in [13].

Acknowledgment. The authors would like to thank anonymous referees for their useful comments. This work is supported by the Grant-in-Aid for Scientific Research (C) (No.22500204), Japan Society for the Promotion of Science.

References

1. Grzymała-Busse, J.W.: Data with Missing Attribute Values: Generalization of Indiscernibility Relation and Rule Induction. In: Peters, J.F., Skowron, A., Grzymała-Busse, J.W., Kostek, B.z., Świniarski, R.W., Szczuka, M.S. (eds.) Transactions on Rough Sets I. LNCS, vol. 3100, pp. 78–95. Springer, Heidelberg (2004)

2. Komorowski, J., Pawlak, Z., Polkowski, L., Skowron, A.: Rough Sets: a tutorial. In: Pal, S., Skowron, A. (eds.) Rough Fuzzy Hybridization, pp. 3–98. Springer (1999)
3. Kryszkiewicz, M.: Rules in incomplete information systems. Information Sciences 113(3-4), 271–292 (1999)
4. Lipski, W.: On semantic issues connected with incomplete information databases. ACM Transactions on Database Systems 4(3), 262–296 (1979)
5. Lipski, W.: On Databases with Incomplete Information. Journal of the ACM 28, 41–70 (1981)
6. Nakamura, A., Tsumoto, S., Tanaka, H., Kobayashi, S.: Rough set theory and its applications. Journal of Japanese Society for Artificial Intelligence 11(2), 209–215 (1996)
7. Okuma, H., Nakata, M., Ślęzak, D., Sakai, H.: Properties on Inclusion Relations and Division Charts in Non-deterministic Information Systems. In: Proc. IEEE GrC Conf., pp. 521–526 (2011)
8. Orłowska, E., Pawlak, Z.: Representation of nondeterministic information. Theoretical Computer Science 29(1-2), 27–39 (1984)
9. Pawlak, Z.: Systemy Informacyjne: Podstawy Teoretyczne. Wydawnictwo Naukowo-Techniczne, Poland (1983) (in Polish)
10. Pawlak, Z.: Rough Sets. Kluwer Academic Publishers (1991)
11. Polkowski, L., Skowron, A. (eds.): Rough Sets in Knowledge Discovery 1: Methodology and Applications. STUDFUZZ, vol. 18. Springer (1998)
12. Polkowski, L., Skowron, A. (eds.): Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems. STUDFUZZ, vol. 19. Physica-Verlag (1998)
13. RNIA software logs, http://www.mns.kyutech.ac.jp/~sakai/RNIA
14. Sakai, H.: Effective procedures for handling possible equivalence relations in non-deterministic information systems. Fundamenta Informaticae 48(4), 343–362 (2001)
15. Sakai, H., Okuma, A.: Basic Algorithms and Tools for Rough Non-deterministic Information Analysis. In: Peters, J.F., Skowron, A., Grzymała-Busse, J.W., Kostek, B.z., Świniarski, R.W., Szczuka, M.S. (eds.) Transactions on Rough Sets I. LNCS, vol. 3100, pp. 209–231. Springer, Heidelberg (2004)
16. Sakai, H., Ishibashi, R., Koba, K., Nakata, M.: Rules and Apriori Algorithm in Non-deterministic Information Systems. In: Peters, J.F., Skowron, A., Rybiński, H. (eds.) Transactions on Rough Sets IX. LNCS, vol. 5390, pp. 328–350. Springer, Heidelberg (2008)
17. Sakai, H., Okuma, H., Nakata, M., Ślęzak, D.: Stable rule extraction and decision making in rough non-deterministic information analysis. International Journal of Hybrid Intelligent Systems 8(1), 41–57 (2011)
18. Tsumoto, S.: Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic. Information Sciences 124(1-4), 125–137 (2000)
19. Ziarko, W.: Variable precision rough set model. Journal of Computer and System Sciences 46(1), 39–59 (1993)

A Method of Internet-Analysis by the Tools of Graph Theory

Sergii V. Kavun, Irina V. Mykhalchuk, Nataliya I. Kalashnykova, and Oleksandr G. Zyma

Abstract. The developed method and its interpretation as a graph model have demonstrated the possibility of the use of optimization methods and further analysis tools in the Internet-analysis. This method has been mathematically formalized and implemented with the aid of linear programming tools. Secondly, the practical implementation of the method has proved that this method has mathematical base and allows a researcher to aggregate statistics [1]. Finally, this analysis can be used as a methodological basis for a preliminary study of the interests of young scientists and graduate students. This may likely provide an objective assessment of the relevance of ongoing or planned research in virtually some limited [1] areas of interest (science's researches, science's activity, and etc.).

1 Introduction

The method of Internet-analysis developed by the authors can be applied in any given area of activity, regardless of its properties and features. The purpose of using this method is to get some assessment or collection of selected concepts (terms), which form the so-called categorical apparatus of scientific research. According to estimates, authors can formulate a conclusion about the need for further research in this area, show the urgency and demonstrated the importance of the calculations, to allocate a narrow major for further research of young scientists. The proposed method can be used to study the activities of prominent scientists, professionals from designated areas of activity. Secondly, this method can be used for demonstrating the importance and the relevance of research, and the conclusions and recommendations drawn, for example, to select consultants (leaders) in their scientific activities.

Sergii V. Kavun · Irina V. Mykhalchuk · Oleksandr G. Zyma
KhNEU, Kharkiv, Ukraine
e-mail: kavserg@gmail.com

Nataliya I. Kalashnykova
UANL, San Nicolás de los Garza, Mexico
e-mail: nkalash2009@gmail.com

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 35–44.
springerlink.com

For a more complete openness, authors will show the possibilities of the research conducted by the authors on the example of the work of scientists in the field of information and economic security, which was selected because of its global relevance. Using a method based on the specifics of the query language that is supported by all search engines and the shape of the query, the results on the set of selected search engines are averaged at a given time interval. Thereby, the dynamism of the study is achieved. The authors showed as an example of some limited areas of interest an interpretation of the developed method with the help of graph theory to be able to use known techniques of optimization and further analysis [6, 10]. Thus, the scope of the developed method of Internet-analysis is multifaceted in its specificity and tolerance in limited areas of interest [6].

2 Problem Description

The description of the problem under consideration has suggested that it's hard to find mathematically defined methods for determining the relevance of the activities, or research of any author(s) or a scientist(s). The solution to this problem in the majority of cases may be too carried out a process that is analytical (with the use a statistical tests or queries) and then enables researches to draw relevant conclusions. This means the decision formulated to overcome the problem doesn't have any mathematical description. This could suggest these types of problem are not easily interpreted (not presented as another known mathematical description) and cannot be formalized in any forms that would allow a researcher to decide any conclusions. In the future a researcher may explore various empirical regularities in the distribution and the relevance of researches in different areas on the same time axis.

The proposed method of Internet-analysis therefore needs to use the elements of the Latent Semantic Analysis (LSA) [4] – namely, the semantic core of investigated area, i.e. keywords and their paradigms. This semantic core (or set $T = \{t_j\}$) can be obtained by using, for example, frequency analysis of publications in selected areas of activity of the author(s). In addition, the proposed method of Internet-analysis may indirectly confirm the known laws of G. K. Zipf [1] (it's the author's suggestion or the hypothesis), but this is the direction for further researches.

3 The Method's Description

The overall description of the method of Internet-analysis is based on the specifics of the query language that is supported by all search engines (for example, Google, Bing, and etc.) and the shapes of the query (with help of this query language). Further to this, the results based on the set of selected search engines are averaged at a given time interval (for example, during one year). The input data for the method of Internet-analysis are (for example and better understanding of work of this method):

1. Set $V = \{v_i\}$, $i = 1 \div 6$, scholars in the field of information and economic security, with their last name, first name and patronymic: Kurkin N.V., Senchagov V.K., Klebanova T.S., Geyets V.M., Oleynikov E.A., Shkarlet S.N.

2. Set T = $\{t_j\}$, j = 1 ÷ 6, categories (terms) of a set of categorical system [5] areas (it's a semantic core [1]): information security (IS) [2], the economic security of the enterprise (ESE), an insider (I) [2], the concept of ESEs (CESE), financial security (FS), the system ESEs (SESE) [5-6, 13]. Categorical apparatus may include (and often and is) a fairly large number of terms, so as the base (to reduce the size of the sample) were selected six concepts, which represent the authors' opinion, some basis for economic security [4-6].
3. Research period: 2001-2011 – set G = $\{g_k\}$, k = 1÷10.
4. Multiple search engines are used, giving a set of results based on average data and using known search servers (for example, Google.com, Yandex.ru, Yahoo.com, Alltheweb.com, Bing.com, AltaVista.com and etc.).

Then V U T = VT ≡ $\{v_i\}$ U $\{t_j\}$ = $\{vt_{ij}\}$, VT = $\{vt_{ij}\}$ – set (or matrix) is averaged which is based on estimates with respect to the scientists selection: a zone of activity in the chosen course of action, the relevance of their research, quantitative saturation of their publications, the relevance of their research, ranking in the world scientific community. The valid (this property was confirmed by a simplified verification using a form of query) results achieved using this method are shown in Table 1.

Table 1 Set VT

T = $\{t_j\}$ V = $\{v_i\}$	IS	ESE	I	CESE	FS	SESE
N.V. Kurkin	659	8	1	0	8	0
V.K. Senchagov	44	21	0	0	94	2
T.S. Klebanova	109	360	0	0	116	1
V.M. Geyets	165	32	0	45	58	126
E.A. Oleynikov	1	9	0	1	13	1
S.N. Shkarlet	2	11	0	1	4	1

Querying search engines in terms of graph theory would be:

$$\{v_i\} \text{ U } \{g_k\} \equiv \text{«Курк*н\_*.В. + 2010» V «Курк*н» + «2010»,}$$

or if using a roman letters for better understanding by foreign reader, this would be

$$\{v_i\} \text{ U } \{g_k\} \equiv \text{«Kurk*n\_*.V. + 2010» V «Kurk*n» + «2010»,}$$

where V is the symbol of operations OR; * – the symbol for the search engine, which means any possible character – this is caused by specific features of spoken language and the translations of attributes of scientists (for example, Куркин (in Russian), Куркін (in Ukrainian)); _ is the space character.

If authors interpret the implementation of the method of online analysis with an aid of graph theory, the authors can construct the corresponding graph and perform the implementation in a mathematical form [2].

Fig. 1 is a Ternary graph of the relationships of scientists attribute (set V) in a given field of study (set T) for a certain time period (set G). This graph model

shows only an interpretation of the parameters used in the method of Internet-analysis. A proof of this method isn't needed, because this method is not exact. This method must only be used for determining trends in selected study area.

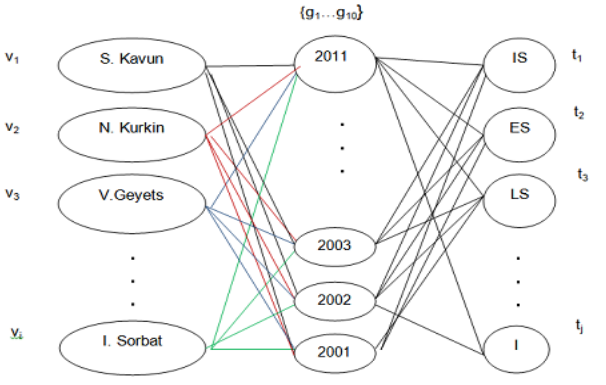


Fig. 1 Ternary graph (a graph model) of the relationships

Each graph has vertices and links between these vertices. The vertices of the security researchers' attributes are on the left side of the graph, whereas the nodes denoting the chosen area of research are positioned on the right side. Coherent link between the two parts of the graph acts as the interim period (starting in 2001 and ending in 2011 [5]).

Weight vg_{ik} edges between vertices v_i and g_k are determined by the average number of links on the results of search engines. Weight vg_{ik} determines the properties of the set $S = \{s_m\}$ with respect to the scientist v_i (urgency, initiative, scale, ambition, etc.) on the set G. The author's proposed the following relationship for calculation a weight vg_{ik} as usual counted as an average of some variables:

$$vg_{ik} = \frac{1}{n}\sum_{k=1}^{x}\{v_i\} \bigcup \{g_k\}, i = const \quad (1)$$

where n is the number of search engines (n = 30);
x is the duration of the study period (x = 10).

Weight of the edge tg_{jk} connecting vertices t_j and g_k are determined by the average number of links on the results of search engines and determines the properties of t_j with respect to g_k. Consequently, authors have the expression

$$tg_{jk} = \frac{1}{n}\sum_{k=1}^{x}\{t_j\} \bigcup \{g_k\}, j = const \quad (2)$$

Equation (2) implements the following request form:

$$\{t_j\} \cup \{g_k\} \equiv \text{«*нсайдер + 2010» V «*нсайдер" + "2010 ",}$$

or if using a roman letters for better understanding by foreign reader, this would be

$$\{t_j\} \cup \{g_k\} \equiv \text{«*nsayder + 2010» V «* nsayder" + "2010 ".}$$

$V = \{v_i\}$ is the set (group) of scientists and specialists conducting research aimed at obtaining a set of properties (S_m) at a given time interval, taking into account the values of t_j. The dimension is determined by an expert.

$T = \{t_j\}$ is the set of categories (terms), defines an area of research and defining its focus and specialization. Dimension is determined by the researcher. Values that make categorical unit area of research are determined by the researcher.

Categorical unit is based on two components: a de-facto (all definitions of scientists, professors, work and studies have been performed in this field) and de-jure (the concept in various dictionaries and standards). The choice is made by a peer review process, in which the selected group of experts-participates. These groups do the job of choice, and, from their point of view, determine the correct answer, which leads to obtaining an acceptable result in the terminology.

$G = \{g_k\}$ is the set of boundary points (years) determined by the specified study period. The dimension is determined by the researcher, or as recommended by the authors from 5 to 10 years. Usually this interval chosen is equal to 1 year, but it's possible to also choose other value.

For further investigation, combine the weight of edge vg_{ik} and weight of edge tg_{jk} in the expression that will determine the aggregate average estimate

$$\frac{1}{n}\sum_{k=1}^{x}\{vg_{ik}\} \bigcup \{tg_{jk}\} = vt_{ij} \tag{3}$$

Equation (3) is implemented in a request form (was showed on an example of research the one from scientists (Table 1) and selected category (I, insider) from semantic core in area of information security):

$$\{vg_{ik}\} \bigcup \{tg_{jk}\} = \text{«Курк*н + *нсайдер+2010» V «Курк*н» + «*нсайдер» + «2010»,}$$

or if using a roman letters for better understanding by foreign reader, this would be

$$\{vg_{ik}\} \bigcup \{tg_{jk}\} = \text{«Kurk*n + *nsayder+2010» V «Kurk*n» + «*nsayder» + «2010».}$$

4 Results

According to the set VT, it is possible to construct graphical dependence, which can be used to analyze [5] the dynamics of change $\{vt_{ij}\}$ (Fig. 2-4).

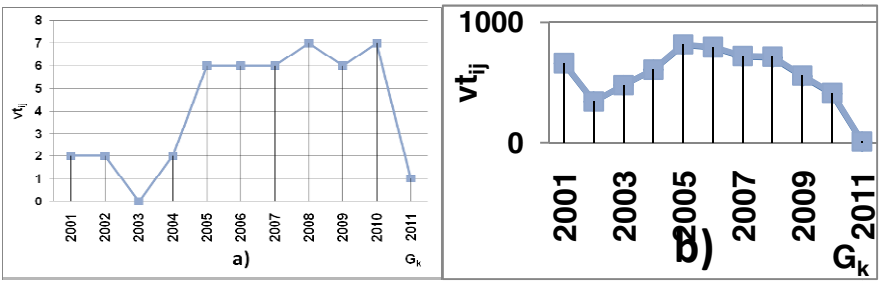


Fig. 2 The results of using a special query (a statistical test) with parameters: a) $v_{i=6}$ = «S.N. Shkarlet», $t_{j=1}$ = «IS»; b) $v_{i=1}$ = «N.V. Kurkin», $t_{j=1}$ = «IS»

For example, on Figure 2(a) are shown the activities of scientist Prof. S. N. Shkarlet [7] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that in period from 2001 to 2005 there has practically been not engaged research in the area of information security. What is suggested is an absence of open publications of this researcher in this area.

On Figure 2(b) are shown the activities of scientist Prof. N. V. Kurkin [8] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that in period from 2002 to 2003 there has been some deduction of science activities of this researcher in the area of information security. What is suggested is a decrease in the number of open publications of this researcher in this area.

On Figure 3(a) are shown the activities of scientist Prof. V. K. Senchagov [9] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that in period from 2006 to 2007 there has been some reduction and after 2010 the researcher ceased to engage in science activities in the area of information security. What is suggested is a decrease in the number of open publications of this researcher in this area.

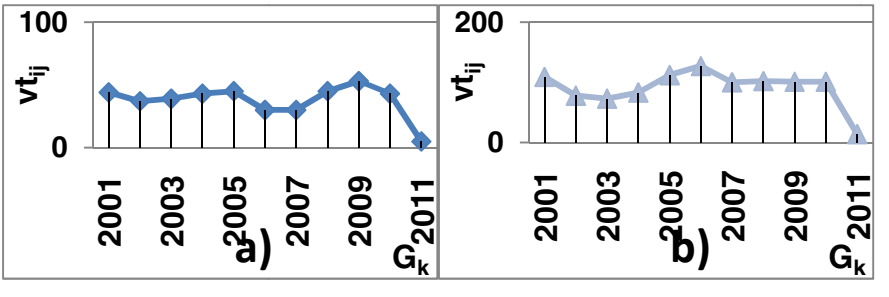


Fig. 3 The results of using a special query (a statistical test) with parameters: a) $v_{i=2}$ = «V.K. Senchagov», $t_{j=1}$ = «IS»; b) $v_{i=3}$ = «T.S. Klebanova», $t_{j=1}$ = «IS»

On Figure 3(b) are shown the activities of scientist Prof. T. S. Klebanova [10] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that in period from 2002 to 2004 there has been some reduction and after 2010 the researcher ceased to engage in science activities in the area of information security. What is suggested is a decrease in the number of open publications of this researcher in this area.

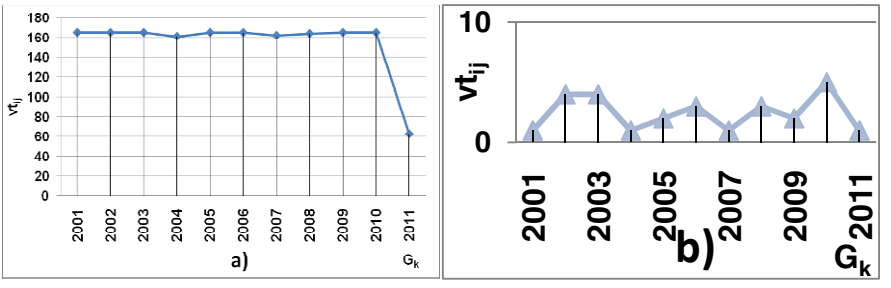


Fig. 4 The results of using a special query (a statistical test) with parameters: a) $v_{i=4}$ = «V.M. Geyets», $t_{j=1}$ = «IS»; b) $v_{i=5}$ = «E.A. Oleynikov», $t_{j=1}$ = «IS»

On Figure 4(a) are shown the activities of scientist Prof. V. M. Geyets [11] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that after 2010 there has been a lower engagement in science activities in the area of information security. What is suggested is a decrease in the number of open publications of this researcher in this area.

On Figure 4(b) are shown the activities of scientist Prof. E. A. Oleynikov [12] on base open publications of this researcher in the area of information security (used a term "IS") for time period from 2001 to 2011. This graph of dynamics has shown that during all research periods, there has been a lower engagement in science activity in the area of information security. What is suggested is a small of number of open publications of this researcher in this area.

The overall trend in population studies of scientists in the field of information security the current year (2011) is very negative, indicating a decline of interest and relevance. However, the finding is subjective, because the time interval of the current year is small and is equal to 3 months [9-7].

5 Discussion

The method of Internet-analysis may seem more straightforward than other mathematical methods of assessment of activity in different research areas. The focus of this mathematical process is on human based analysis and derived conclusions to solve problems in research areas (it's a choice of: semantic core, area of researches, author as a person and etc.) whilst remaining relevant to assist researchers. There

are certain additions to our ternary graph (this Ternary graph can have greater dimension) which seemed likely to result in further improvement [2, 3].

One of the most important moments is the ability to identify keywords and them paradigms in investigated area. Given that there is not a large quantity of keywords to be used in the paradigms in semantic core of investigated area, such a situation may have consequences. Thus what might become possible is to receive true results with the help of the proposed method. One of the consequences is associated with, is the use of near identical keywords and how they are used within the different paradigms in semantic core of the proposed method. Such a result, may lead to false or very subjective results.

Despite the gap between human and automated performance, the current level of accuracy of the method of Internet-analysis approach has already led to some promising early results in research of activity of some scientists. For example, in a separate study of activity of authors Prof. A. B. Nemeth and S. Z. Nemeth in the book EUCLIDEAN VECTOR LATTICES in 2011 (Springer Mathematics Review), authors used this method to track across multiple time periods (this multiple time periods showed on Figure 5). Each trial consisted of a single query by semantic core on increasing number of search servers. In this separate study, broad learning effects identified by the queries were consistent with effects found using manually-scored pre- and post-queries. Our automated method of Internet-analysis revealed a previously unknown interaction between time spacing, the proportion of supportive queries on base semantic core, and activity of some scientists [14-15].

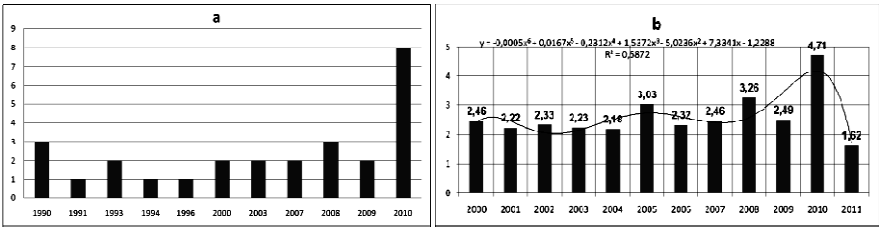


Fig. 5 The results of research of science activity with help of proposed method: a) distribution of numbers of links (the open science works) in selected time period; b) distribution of numbers of links of the identify keywords (from semantic core) in investigated area

With some adjustments, the same method of Internet-analysis described in this paper may be applied to the problem of finding or determining a person's (for example, insiders [2, 6, 3]) activity on the Internet.

6 Conclusions

Authors have presented results of using the method of Internet-analysis, based on examples of the definition of professional activity of different scientists. The obtained results are likely to allow users to receive an objective assessment of the

significance, importance and necessary of research, some scientists or any other specialist with different focus areas including enterprises. The Authors described forms for query as examples using a query language that can be used on any search servers. Authors showed on example (Springer Mathematics Review) that can be to use this method of Internet-analysis that can assist in writing grounded reviews. This method of Internet-analysis was based on known query languages, which can be used on complex problems with other tools or applications.

The developed method and its interpretation as a graph model have demonstrated the possibility of the use of optimization methods and further analysis tools.

In addition, expressions (1)-(3) can mathematically formalize the implementation of the method of Internet-analysis, which will further lead to a linear programming problem. The practical implementation of the method has suggested that this method has a solid mathematical base and allows one to aggregate statistics.

Finally, this analysis can be used as a methodological basis for a preliminary study of the interests of young scientists and graduate students. This will get (to a certain extent) an objective assessment of the relevance of ongoing or planned research in virtually some limited areas of interest (for example, science's researches, science's activity, choice of mentor, study of people's activity, search and detection of criminals, and etc.).

References

1. Zipf, G.K.: Human Behavior and the Principle of Least Effort. Addison-Wesley, Cambridge (1949)
2. Kavun, S.: Informational security: textbook / Kavun, S., vol. 368. Pub. of INZHEK, Kharkiv (2009)
3. Ponomarenko, V.: Conceptual foundations of economic security: monograph / Ponomarenko, V., Kavun, S., vol. 256. Pub. INZHEK, Kharkiv (2008)
4. Landauer, T.K., Foltz, P.W., Laham, D.: An Introduction to Latent Semantic Analysis. Discourse Processes 25, 259–284 (1998)
5. Kavun, S.: Analysis of categorical apparatus in the field of economic and information security / Kavun S, Mikhalchuk I / Economic of development: Science magazine, vol. 3 (51), pp. 9–14. Pub. INZHEK, Kharkiv (2009)
6. Kavun, S.: Model of insiders in the company. In: Kavun, S., Sorbat, I. (eds.) Securitatea Informațională: ed. a 5-a Jubiliară Conf. Intern., April 16-17, vol. 76, ASEM, Chişinău (2008)
7. Shkarlet, S.: The economic security of enterprise: innovation aspect: monograph / Shkarlet, S., vol. 436. Books view of the NAU, Kiev (2007)
8. Kurkin, N.: Managing the economic security of the enterprise: monograph/ Kurkin, N., vol. 452. Art-Press, Dnepropetrovsk (2004)
9. Senchagov, V.: Strategic objectives mechanism to ensure the economic security. In: Senchagov, V.: Problems of Theory and Practice of Management, vol. (3), pp. 18–23 (2009)
10. Ponomarenko, V., Klebanova, T., Chernov, N.: Economic security in the region: analysis, evaluation, and prediction: monograph, vol. 143. INZHEK, Kh (2004)

11. Geyets, V.: The modeling of economic security: power, region, enterprise: monograph / Geyets, V., Kizim, M., Klebanova, T., Chernyak, O., vol. 240. Pub. INZHEK, Kharkiv (2006)
12. Oleynikov, E.: Principles of Economic Security (State, region, company, person. In: Oleynikov, E. (ed.) Moscow: JSC "Business School" Intelligence-Synthesis, vol. 288 (1997)
13. Kavun, S.: Method for constructing a security policy organization. In: Kavun, S., Shubina, G.: Business-Inform., vol. (1-2), pp. 96–102 (2005)
14. Murdoch, C.: Economic factors as objects of security: Economics security & vulnerability. In: Knorr, K., Trager, F.: Economics Interests & National Security, vol. 867, Lawrence (2001)
15. Eric Johnson, M.: Managing Information Risk and the Economics of Security, 1st edn. Interperiodica distributed exclusively, vol. 347. Springer Science + Business Media LLC (2008)

A Multi-agent Decision Support System for Path Planning under Reciprocal Constraints

Miri Weiss-Cohen, Gideon Avigad, and Erella Eisenstadt

Abstract. This paper formulizes and solves a special class of trajectory optimization problems focusing on the optimization of different kind of agents which are coupled through reciprocal constraints. Searching for a solution to the problem at hand may not only stipulate planning a longer path for each agent, but also choosing slower travel speeds in order to coordinate multiple paths. A novel approach is formulated which uses the representation of a path by a telescopic antenna. The length and orientation of the antenna's sections are evolved by a genetic algorithm. The approach is demonstrated and studied through solving and simulating several path-planning problems.

1 Introduction

Over the last three decades researchers have extensively investigated the design, control, and planning of paths/trajectories, with the aim of optimizing the movement with respect to a single objective (e.g., [1]) or multiple objectives (e.g.,[2]). The first studies and implementations considered the optimization of the paths of a single agent. Today researchers are focusing on optimizing sets of cooperating agents. Planning paths/trajectories for several agents is significantly more difficult than the path-planning problem for single agent systems, since the size of the joint state space of each agent increases exponentially with the number of agents [3]. Researchers and practitioners have recognized that when several agents cooperate, many tasks can be performed more efficiently and robustly [4]. Examples of tasks are exploration [5], search and rescue [6], and transportation of large objects [7].

Miri Weiss-Cohen
Software Engineering Department –ORT, Braude College of Engineering

Gideon Avigad · Erella Eisenstadt
Mechanical Engineering Department –ORT, Braude College of Engineering
e-mail: {miri,gideona,erella}@braude.ac.il

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 45–54.
springerlink.com

The current paper deals with a wider definition of the coupled trajectories problem. Here the task of the agents is not mutual; each agent has its own task and therefore its own related optimization problem, possibly operating in totally different workspaces. Nevertheless, these individual optimization problems are coupled through reciprocal constraints.

The methods for solving coupled trajectory problems may be divided into two categories [9]. In the *centralized* approach, the configuration spaces of the individual agents are combined into one composite configuration space, which is then searched for a path that is appropriate for the whole composite system [10,11]. In contrast, the *decoupled* approach first computes separate paths for the individual agents and then resolves possible conflicts of the generated paths [12-13].

Genetic Algorithms (GAs) are used extensively for solving single objective problems and are also an appealing option when Multi-Objective Optimization Problems (MOPs) are to be solved. In the case of contradictory objectives, no universally accepted definition of "optimum" exists for MOP solutions, as it does for single-objective optimization problems [14]. Searching a multi-objective design space for optimal solutions by Evolutionary Computation EC approaches is commonly referred to as Evolutionary Multi-objective Optimization (EMO) [8]. A Multi-Objective Evolutionary Algorithm (MOEA) is an EMO algorithm that searches for a solution in a multi-criteria space using some inspiration from evolutionary theories.

2 Motivation and Approach

The motivations for the current work are: (1) to introduce and formulate a class of problems for which several paths/trajectories should be optimized while meeting reciprocal constraints, and (2) to highlight the multi-objective aspects of such problems and the need to introduce tradeoffs while solving them (Decision Support System).

A common path optimization problem may be defined as follows:

Find a path P(x,t) in order to minimize $\Psi(\mathrm{P})$ subject to $g(x,t) \le 0$ and $h(x,t) = 0$, where x is a vector of the position coordinates and t is the time. $\Psi(P)$ may be the path length, path time, change of direction within the path, etc.

The current problem is defined as follows: Find paths (a solution) $P(x,t)$ where $P(x,t) = [P^1(x_1,t),.....,P^K(x_K,t)]^T$ so that:

$$Min(\Psi(P)),\ \Psi(P) = [\Psi^1(P^1(x_1,t),....\Psi^i(P^i(x_i,t),.....,\Psi^K(P^K(x_K,t)]^T \qquad (1)$$

subject to: $g_j(P(x,t)) \le 0$ for $1 \le j \le J$ and $h_e(P(x,t)) = 0$ for $1 \le e \le E$, where x_i is the position coordinate of the i-th path of x and t is the time. $\Psi^i(P^i(x_i,t)$ is the objective of the i-th path's planning, which may differ from one path to

another. $g_j(P(x,t))$ and $h_e(P(x,t))$ are inequality and equality constraints, respectively. The dependence of the constraint functions (g and h) on several paths engenders coupling between the paths.

Two approaches are considered to search for the solution to the problem at hand (Equation 1). In the first approach, which we refer to as the posed single objective problem, the optimization problem is defined as follows:

$$Min(\Phi(P)),\ \Phi(P) = \sum_{i=1}^{K} \lambda_i \Psi^i(P^i(x_i,t)) \tag{2}$$

subject to: $g_i(P(x,t)) \le 0$ for $1 \le i \le K$, $r_j(P(x,t)) \le 0$ for $1 \le j \le J$

where g_i are the specific constraints of the i-th paths and r_j are the reciprocal constraints of the j-th paths. λ_i is the i-th trajectory weight, which reflects the designer's view concerning the importance of this path with respect to the other paths. This weight may also include a scalarizing element. Here, we utilize λ just for scalarizing, as we regard all problems to be equally important. Note that the definition in Equation (2) is a weighted sum approach. Previous research has shown that this method often produces poorly distributed solutions along a Pareto front [15], and that it does not find Pareto optimal solutions in non-convex regions. Yet, here this weighted sum will not be used to find the Pareto set by altering weights but rather through a sequential process. In using the following as a Decision Support System, tradeoffs may be represented by extending the posed single objective optimization approach. In some cases this may lead to an approach that resembles the sequential constrained optimization procedure suggested by [8].

3 The Revolute Prismatic Chain (RPC) Algorithm

The location and speed of each path/trajectory must be determined, leaving both as design variables. The need to change location and speed in one path due to changes in another path while avoiding obstacles is not straightforward or easy. A new approach, termed the Revolute Prismatic Chain (RPC), gradually extracts (prismatic) and tilts (revolute) the problem-defined parameters by evolutionary methods. The RPC is constructed from a predefined number of sections. A section represents the distance the agent travels at each time step. Each section is anchored at the end of the preceding section (the first section is anchored to the starting point). The angle at which each section is extended is coded as a parameter, as is the length of each section. This means that the speed of travel is coded by this length (length per time is the speed of travel). The maximal length is the maximal speed the agent may reach. Note that the speed at each section is the mean velocity at this section. Figure 1 depicts a five-section RPC. Note that the end of the RPC is not anchored to the end point and is located wherever the last section ends.

In the current study, we utilized an elite-based EA in order to search for the solutions to our problems. An individual within the population is constructed out of several chromosomes, each coding one trajectory out of the trajectories of the coupled problems.

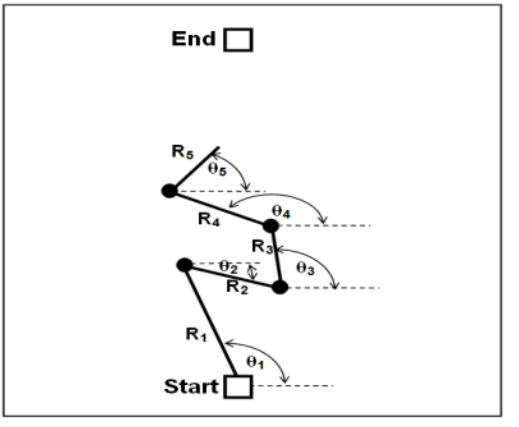


Fig. 1 The RPC with five sections. The length of each section and its angle with respect to the previous section are coded with the evolutionary algorithm code.

The coded parameters are decoded into intervals such that $-180^{o} \le \theta_i \le 180^{o}\ ; 0 \le R_i \le 1\ \ foralli = 1,.....N$. The number of RPC sections, N, and the number of members in a set (number of RPCs), K, are predefined. The influence of the constraint violations on the evolution was introduced into the algorithm through a procedure detailed in the following code. The EA is an elite GA-based algorithm that is given by the following pseudo code:

1. Initialize a population P_t with n individuals. Also, create a population $Q_t = P_t$
2. Combine parent and offspring populations and create $R_t = P_t \cup Q_t$.
3. For each individual x in R_t compute: Φ(x), d(x), δ(x) as detailed in calculations I, II, III respectively.
4. Create Elite population P_{t+1} of size n from R_t (procedure I)
5. Create offspring population Q_{t+1}^{*} of size n from R_t by Tournament selection (procedure II)
6. Perform Crossover and Mutation to obtain Q_{t+1}^{**} from Q_{t+1}^{*}.

Calculation I: Φ(x)

1. Find trajectories: $P(x,t)=[P^1(x_1,t),.....,P^K(x_K,t)]^T$
2. Find $\Psi(P(x,t))=[\Psi^1(P^1(x_1,t),....\Psi^i(P^i(x_i,t),.....,\Psi^K(P^K(x_K,t)]^T$, and compute $\Phi(x)=\sum_{i=1}^{K}\lambda_i\Psi^i(P^i(x_i,t)$

Calculation II: d(x)

1. Find the distance from the i-th RPC end point to the i-th end point as follows

$$d(x)=\begin{cases}\neg\exists g_i(x_i,t)\mid g_i(x_i,t)>0 \quad Then \ \max\limits_{1\le i\le K}\left|x_{end}-P(x_i,t_{final})\right| \\ \exists g_i(x_i,t^*)\mid g_i(x_i,t^*)>0 \quad Then \max\limits_{1\le i\le K}\left|x_{end}-P(x_i,t^*)\right|\end{cases}$$

Calculation III: δ(x)

1. For all trajectories of x compute
$$\delta_{i,l}=\left|P^i(x_i,t)-P^l(x_l,t)\right| \ for\ all\ i=1,....,K;\ i\neq l$$
2. Compute $\delta(x)=\Gamma(\delta_{i,l})$ where Γ, is a problem related function

Note: In all the examples, $\Psi^i(P(x_i,t))=Length\,of\,the\,trajectory\,P(x_i,t)$, $\lambda_i=1\ for\,all\,i=1,...,K$. In the first example the reciprocal constraint is to keep the distance between trajectories smaller than d_{min}, $\delta(x)=\overline{(\delta_{i,l})}$; in all other examples, the reciprocal constraint is to achieve a meeting point, $\delta(x)=\min\limits_{\forall i,l}(\delta_{i,l})$.

4 Examples

In this section, two examples are solved. The examples utilize a 50% crossover with Gaussian distributed mutation with probability of 5%. We note that such rates are common in GAs [19]. The first example involves two agents that start in different locations and have to reach their respective target points by traveling the shortest distance while avoiding obstacles. This means that $\Phi=[L_1,L_2]^T$, where L_1,L_2 are the lengths of the two paths. The reciprocal constraint is that the distance between the agents is not to exceed a predefined distance. In terms of Equation 1, the problem may be defined as follows:

$$\Phi=[L_1,L_2]^T\ \forall t\neg\exists P^i(x_i,t),P^j(x_j,t)\in P(x,t):\left|P^i(x_i,t)-P^j(x_j,t)\right|\ge\varepsilon_1$$

(keeping the distance amongagents at all times)

$$\forall P^i(x_i,t)\in P(x,t)\wedge j\in J:\neg\exists P^i(x_i,t):\left|P(x_i,t)-O_j(x,t)\right|-\varepsilon_2\le 0$$

(obstacle avoidance),

where L_1 and L_2 are the travel distances of the two agents, $O_j(x,t)$ is the coordinates of the j-th obstacle and $\varepsilon_1, \varepsilon_1$ are predefined distances.

Figure 2 a-b depicts the 300 and 400 generations, respectively. Figures 3a and 3b, respectively, depict the solutions after running the same algorithm but with an RPC with 15 sections after 200 and 400 generations.

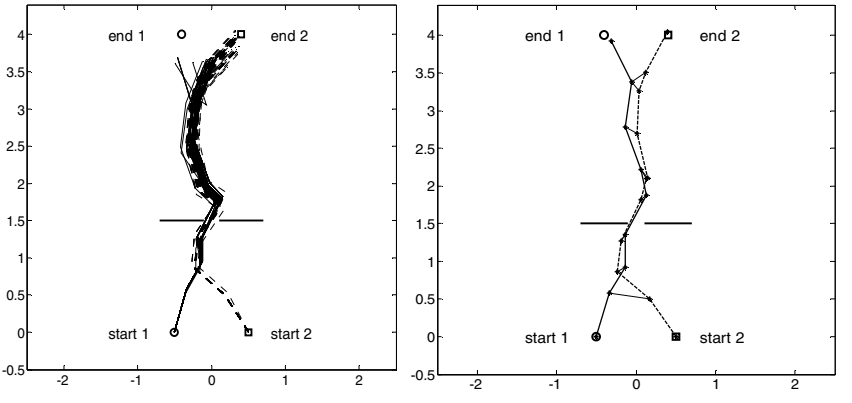


Fig. 2 a-b The evolution of a 10-section RPC; generation 300 (2a) and final generation (2b)

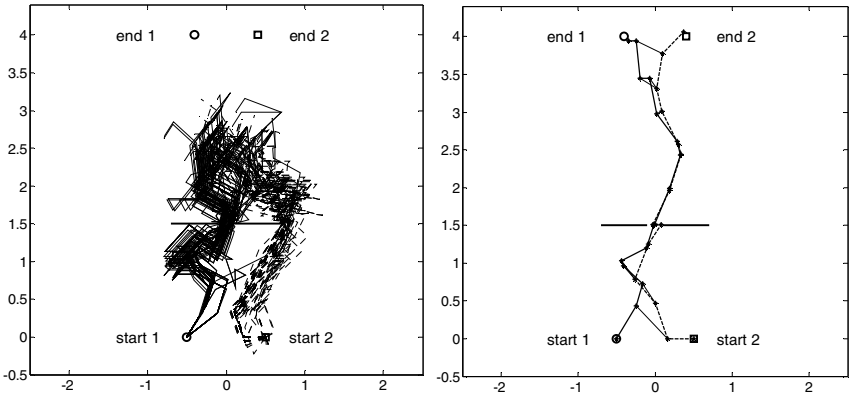


Fig. 3 a-b The evolution of a 15-section RPC; generation 200 (3a), final generation (3b).

Figures 4 a-b show the series of solutions that results from repeating the algorithm run by following the sequential optimization-constrained procedure of the results of Section 3 in a series of solutions. Note that for each run, new obstacles are positioned (automatically) along the already found trajectory. The added obstacles are designated by short bold lines. The newly added obstacles may be bigger than those we chose to use in the current example, constraining the

subsequent paths to be located further away from their former ones. The results clearly show how new paths are found.

A plot of all obtained solutions on the same plot is shown in Figure 5a. As a last step in the sequential optimization-constrained procedure, the performances of each solution (related path lengths of each member) are plotted in the objective space, and the Pareto front is identified. For the current problem, this is depicted in Figure 5b, where the Pareto set appears to include solutions 1, 2 and 6.

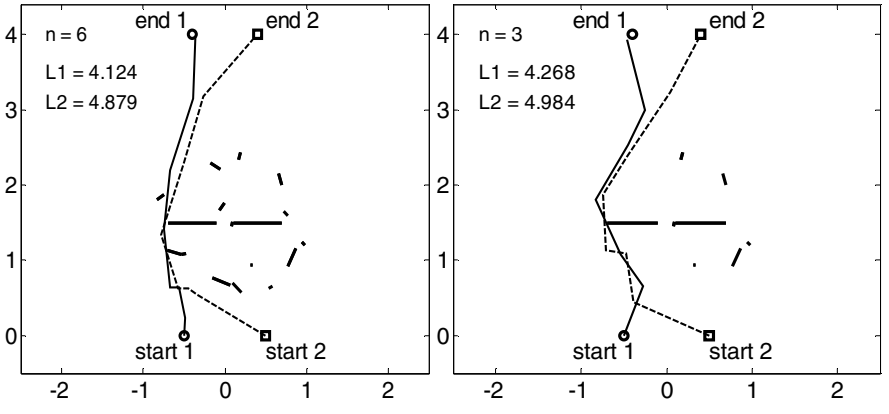


Fig. 4 a-b Adding constraints (obstacles) once a solution is obtained. The constraints are new obstacles set along the previously found trajectory.

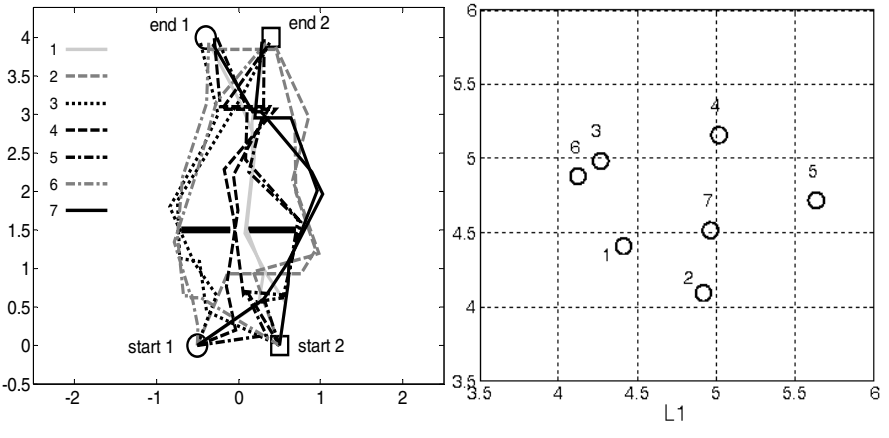


Fig. 5a Result of the sequential run.

Fig. 5b Plotting of pairs L_1 and L_2

The purpose of second example is to discuss some statistical issues related to the problem at hand so as to try to understand why an evolutionary multi-objective optimization fails to find the Pareto set. In this example, a map free of obstacles is

considered. The reciprocal constraint is the need to find a meeting point somewhere along the agents' paths. Figures 6a and 6b depict the initial and final generation of the evolutionary search. At the meeting point, the distance between the agents is smaller than the maximal constraint d_{min}. Running the algorithm 100 times and replotting the meeting point on the same plot, as depicted in Figure 7a, reveals that most of the points are concentrated around the center of the enclosed. This statistical data is represented as box plots in Figure 7b.

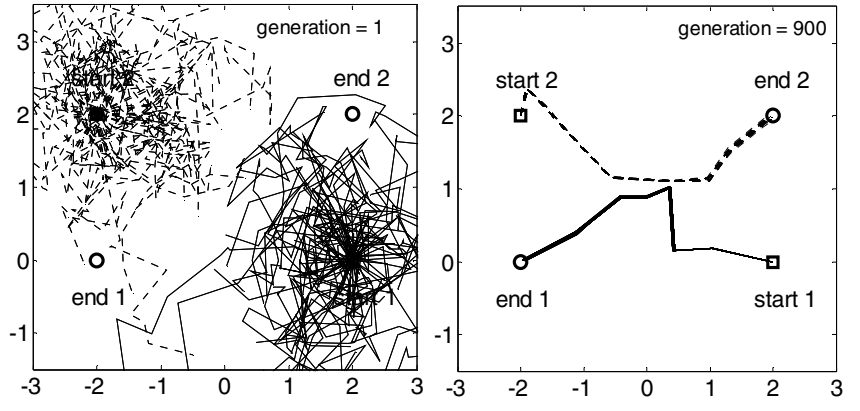


Fig. 6a-b The populations of pairs of trajectories (one member is designated by a bold curve while the other by a dashed curve) in the first generation (19a) and in final generation (19b).

The averages of the agents' path lengths (L_1 and L_2) are almost the same, indicating that no preference is given to one agent over the other. This is conceivable because the weights (λ_1, λ_2) of the paths within the utility function, Φ, are equal (both equal 1.0).

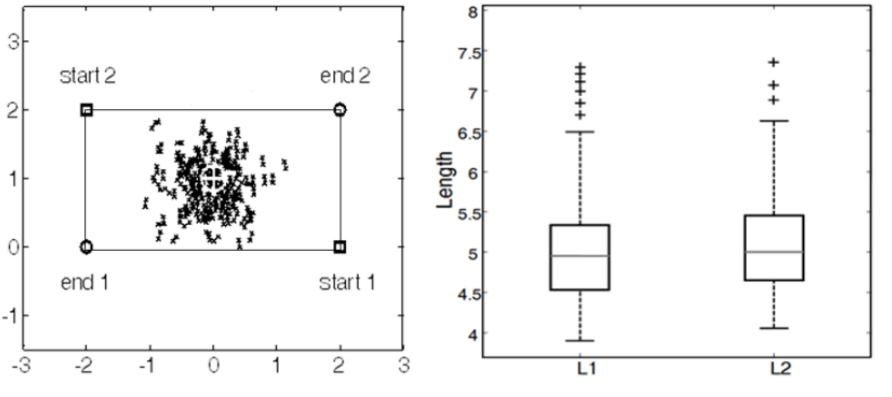


Fig. 7a-b Meeting points (designated by x signs) which were found after 100 different runs (7a) and related statistic data (7b).

The smoothening procedure is implemented on the solution shown in Figure 8b. A major improvement has been achieved, as depicted in Figure 8a. The final generational optimal solution is designated by the dashed curves, while the mended trajectories are designated by the bold curves. Clearly, the paths are shorter, while the meeting point is maintained. This improvement of the results is consistent with respect to all the solutions found in the 100 runs, as can be observed from the box plots in Figure 8b. The path lengths have been shortened (from around 4.8 in average to an average of around 4.4).

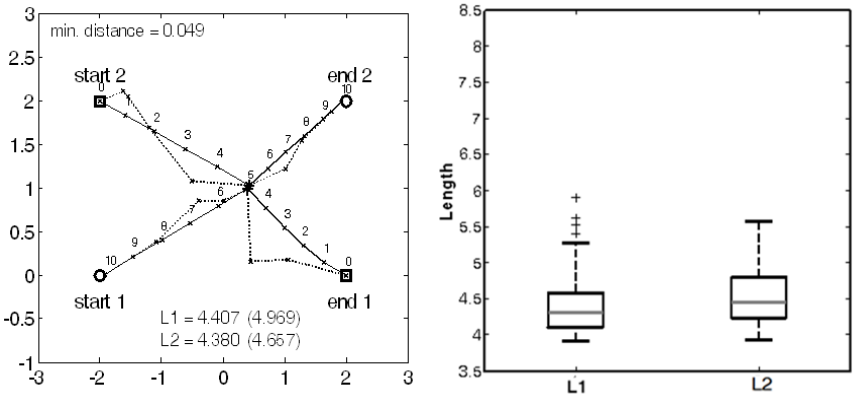


Fig. 8 A solution found on one of the runs and its corresponding mended solution (dashed and bold trajectories respectively in 8a) and the statistical data representation (8b).

In Figure 9 the initially evolved solutions are plotted on the same objective space together with the mended solutions (smoothed by the hybrid algorithm). The solution performances initially found are designated by gray circles and the mended solution performances by black circles. A gray circle may become a black one, not by the hybrid approach but by running the evolutionary algorithm for a greater number of generations.

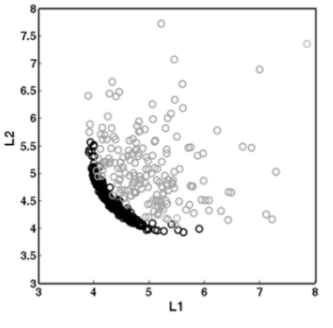


Fig. 9 Lengths of trajectories found in 100 different runs (designated by a gray circle for each pair), and these lengths after smoothing by the a posteriori executed mending algorithm.

5 Conclusions

A novel algorithmic approach has been proposed, in which a path is represented by a telescopic antenna. The length and orientation of the antenna's sections are evolved by a genetic algorithm. The length is limited by the maximal speed of the UMV. The proposed algorithm is tested by solving several coupled trajectory planning problems. The results show that the approach is easily adapted to different problem settings and can solve a range of problems. In the study, several algorithmic approaches were considered. These included a sequential approach and a simultaneous one. The latter, although intuitively better in representing the tradeoff, was discussed.

References

[1] Cong, Y.Z., Ponnambalam, S.G.: Mobile Robot Path Planning using Ant Colony Optimization. In: 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Suntec Convention and Exhibition Center, Singapore, July 14-17 (2009)

[2] Castillo, O., Trujillo, L., Melin, P.: Multiple objective genetic algorithms for path-planning optimization in autonomous mobile robots. Soft Computing 11, 269–279 (2007)

[3] Bennewitz, M., Burgard, W., Thrun, S.: Optimizing schedules for prioritized path planning of multi-robot systems. In: Proc. IEEE ICRA, pp. 271–276 (2001)

[4] Donald, B., Gariepy, L., Rus, D.: Distributed manipulation of multiple objects using ropes. In: Proc. IEEE ICRA, pp. 450–457 (2000)

[5] Burgard, W., Moors, M., Fox, D., Simmons, R., Thrun, S.: Collaborative multi-robot exploration. In: Proc. IEEE ICRA, pp. 476–481 (2000)

[6] Jennings, J.S., Whelan, G., Evans, W.F.: Cooperative search and rescue with a team of mobile robots. In: Proc. IEEE ICRA, pp. 193–200 (1997)

[7] Sugar, T., Kumar, V.: Control and coordination of multiple mobile robots in manipulation and material handling tasks. In: Corke, P., Trevelyan, J. (eds.) Experimental Robotics VI. LNCIS, vol. 250. Springer, Berlin (2000)

[8] Avigad, G.: A simultaneous EMO for the solution of the multi-multi-objective optimization problem. In: IEEE CEC 2007, pp. 2117–2124 (2007)

[9] Latombe, J.C.: Robot motion planning. Kluwer Academic Publishers, Boston (1991)

[10] Azarm, K., Schmidt, G.: A decentralized approach for the conflict-free motion of multiple mobile robots. In: Proc. IEEE/RSJ IROS, pp. 1667–1674 (1996)

[11] Kavraki, L., Svestka, P., Latombe, J.C., Overmars, M.: Probabilistic road maps for path planning in high dimensional configuration spaces. In: Proc. IEEE ICRA, vol. 12(4), pp. 566–580 (1996)

[12] Warren, C.: Multiple robot path coordination using artificial potential fields. In: Proc. IEEE ICRA, pp. 500–505 (1990)

[13] Kaminer, I., Yakimenko, O., Pascoal, A., Ghabcheloo, R.: Path generation, path following and coordinated control for time-critical missions of multiple UAVs. In: Proc. Amer. Control Conf., Minneapolis, Minnesota, USA, June 14-16 (2006)

[14] Parker, L.E.: Path planning and motion coordination in multiple mobile robot reams. In: Meyers, R.A. (ed.) Encyclopedia of Complexity and System Science. Springer (2009)

[15] Pareto, V.: Manual of political economy. The MacMillan Press (original edition in French) (1927/1971)

A New Way to Make Decisions with Paired Comparisons

Youichi Iida

Abstract. In this paper we propose a decision-making method with paired comparisons, which calculates a ratio of evaluation values of alternatives with respect to a goal. If we have concrete evaluation values for each alternative with respect to criteria, the ratio calculated is the same as one with weighted summation method. In this paper we show the procedure of this method and its validity with mathematics. Furthermore, we propose a way to check validity of the final answer obtained with this method. We don't discuss what kind of paired comparisons we should utilize, while we use paired comparisons in the Analytic Hierarchy Process (AHP) and the Analytic Network Process by T.L.Saaty in an example.

1 Introduction

Weighted summation method is a traditional and useful multi-criteria decision-making method. On the other hand, we sometimes need only ratio of evaluation values of alternatives with respect to a goal, for example, the case we need ranking among alternatives. In this paper, we propose a decision-making method using the ratio of evaluation values of alternatives with respect to a goal. When we have concrete evaluation value of each alternative with respect to criteria, the ratio obtained with this method is the same as one with weighted summation method.

This method is using lots of paired comparisons, though they were hardly dealt with in this paper. We use paired comparisons in the Analytic Hierarchy Process (AHP) and the Analytic Network Process (ANP) proposed by T.L.Saaty in order to show the procedure of the proposed method in Section 3. This method is like AHP and ANP (cf. [1, 2]).

We show the purpose of the method by an example in Section 2 and explain the procedure of the method in Section 3. Furthermore, we explain mathematically

Youichi Iida
Tokyo University of Science, Suwa, Nagano, Japan
e-mail: iida@rs.suwa.tus.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 55–66.
springerlink.com

its validity in Section 4 and propose how to check validity of the final answer in Section 5. Finally we have conclusions in Section 6.

2 The Purpose of the Method Proposed

We review the procedure of weighted summation method by a simple example in order to show the purpose of the method. We consider a situation where someone is about to choose one from three fruit baskets for himself/herself, which are Basket 1, Basket 2 and Basket 3. These baskets contain some apples, oranges, peaches and bananas as in Table 1. Here we suppose that his/her preferences among four kinds of fruits are in Table 2, the sum of which is 1.

Table 1 The number of fruits in baskets

	Apples	Oranges	Peaches	Bananas
Basket 1	1	4	2	5
Basket 2	2	5	3	1
Basket 3	4	3	5	2

Table 2 Preferences among fruits in baskets

	Apples	Oranges	Peaches	Bananas	Total
Preferences	0.542	0.247	0.134	0.077	1

Then we have from Tables 1 and 2

$$\begin{pmatrix} \text{Preference of Basket 1} \\ \text{Preference of Basket 2} \\ \text{Preference of Basket 3} \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 5 \\ 2 & 5 & 3 & 1 \\ 4 & 3 & 5 & 2 \end{pmatrix} \begin{pmatrix} 0.542 \\ 0.247 \\ 0.134 \\ 0.077 \end{pmatrix} = \begin{pmatrix} 2.183 \\ 2.798 \\ 3.733 \end{pmatrix}.$$

Consequently, we have Basket 1 (2.183) $\prec$ Basket 2 (2.798) $\prec$ Basket 3 (3.733) and so he/she chooses Basket 3 out of three baskets. This is weighted summation method. By the way, in this example the ranking among three baskets is essential for him/her. Namely, if he/she has the continued ratio Basket 1 : Basket 2 : Basket 3$(=2.183 : 2.798 : 3.733) = 0.251 : 0.321 : 0.428$ and the ranking Basket 1 (0.251) $\prec$ Basket 2 (0.321) $\prec$ Basket 3 (0.428), then he/she chooses Basket 3 as well.

The purpose of the method proposed in this paper is to obtain the ratio of priorities among alternatives with respect to a goal when we don't know concrete evaluation values of alternatives with respect to criteria like Table 1 beforehand.

3 The Procedure of the Method Proposed

The procedure of the method is the following:

(Step 1) Construct a hierarchy to represent a given problem. In this paper we deal only with three-level hierarchy with one goal, criteria and alternatives for the simplicity.

(Step 2) Calculate weights of criteria with respect to the goal with paired comparisons. We use ones in the AHP/ANP through this paper.

(Step 3) Calculate the evaluation ratio table of alternatives with respect to criteria under the goal.

- **(Step 3-1)** Calculate weights of alternatives with respect to criteria with paired comparisons.
- **(Step 3-2)** Calculate weights of criteria in each alternatives under the goal with paired comparisons. These weights are the rates that criteria occupy in each alternative.
- **(Step 3-3)** Adjust weights obtained in (Step 3-1) with weights in (Step 3-2). We obtain a number of tables which are the same as the number of alternatives.
- **(Step 3-4)** Synthesize all tables in Step (3-3) to one evaluation ratio table of alternatives with respect to criteria under the goal.

(Step 4) Combine weights of criteria in (Step 2) and weights of alternatives in (Step 3), i.e., (Step 3-4), to obtain the ratio of evaluation values of alternatives with respect to the goal.

(Step 5) Check validity of final answer in (Step 4). This is due to validity of the table obtained in (Step 3-4). We deal with this step in Section 5.

Now we explain the procedure using concrete numbers. We use SuperDecisions 2.0.8, which is famous software for the AHP/ANP, in calculating weights of criteria and alternatives and C.I., and Microsoft Office Excel 2007 in adjusting, synthesizing and combining weights of alternatives.

(Step 1) We represent a given problem with a hierarchy of one goal, criteria, subcriteria and alternatives. In this explanation we suppose the three-level hierarchy with four criteria C_1, C_2, C_3 and C_4 and three alternative a_1, a_2 and a_3 as in Figure 1.

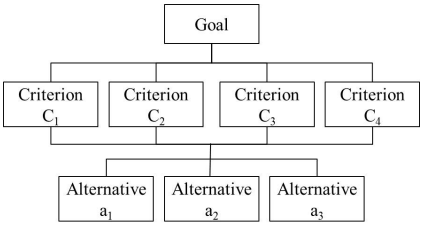


Fig. 1 A three-level hierarchy

(Step 2) We calculate weights of criteria with respect to the goal. Here we use paired comparisons with $1, 2, \ldots, 9, 1/2, \cdots, 1/8$ and $1/9$ in the AHP and obtain Table 3. Then we use , for example, the following question: Which is better between two criteria with respect to the goal? and how many times better?

Table 3 Paired comparison matrix of criteria with respect to Goal

Goal	C_1	C_2	C_3	C_4	Weights
C_1	1	5	7	9	0.657
C_2	1/5	1	3	5	0.203
C_3	1/7	1/3	1	3	0.094
C_4	1/9	1/5	1/3	1	0.046
					C.I. = 0.064

Consequently, we have $C_1 : C_2 : C_3 : C_4 = 0.657 : 0.203 : 0.094 : 0.046$. These weights are normalized, i.e., the sum of them is 1.

(Step 3-1) We calculate weights of alternatives with respect to criteria with paired comparisons. Then we use, for example, the following question: Which is better between two alternatives with respect to a criterion? and how many times better? We obtain Table 4 with paired comparisons in the AHP and Table 5 as a result. The weights in these tables are normalized in the AHP, although this normalization isn't needed in the method proposed.

Table 4 Paired comparison matrix among alternatives with respect to criteria

C_1	a_1	a_2	a_3	Weights
a_1	1	3	1/3	0.259
a_2	1/3	1	1/5	0.105
a_3	3	5	1	0.637
				C.I. = 0.037

C_2	a_1	a_2	a_3	Weights
a_1	1	1/7	1/2	0.103
a_2	7	1	3	0.682
a_3	2	1/3	1	0.216
				C.I.=0.003

C_3	a_1	a_2	a_3	Weights
a_1	1	3	2	0.540
a_2	1/3	1	1/2	0.163
a_3	1/2	2	1	0.297
				C.I. = 0.009

C_4	a_1	a_2	a_3	Weights
a_1	1	1/3	1/5	0.109
a_2	3	1	1/2	0.309
a_3	5	2	1	0.582
				C.I.=0.004

(Step 3-2) We calculate weights of criteria in each alternative like the ANP. Firstly, we take Alternative a_1, that we call a basis and denote with $\tilde{a_1}$. Then a decision maker compares criteria in a_1 using, for example, the following question: Which of two criteria occupies Alternative a_1 more? and how many times more? Consequently, we have $C_1 : C_2 : C_3 : C_4 = 0.565 : 0.262 : 0.118 : 0.055$ for $\tilde{a_1}$ (see the left table in Table 6). Similarly, we have $C_1 : C_2 : C_3 : C_4 = 0.111 : 0.732 : 0.049 : 0.108$ for $\tilde{a_2}$ and $C_1 : C_2 : C_3 : C_4 = 0.596 : 0.266 : 0.042 : 0.097$ for $\tilde{a_3}$ (Table 6). These

Table 5 Weights of alternatives with respect to criteria

Goal	C_1	C_2	C_3	C_4
a_1	0.259	0.103	0.540	0.109
a_2	0.105	0.682	0.163	0.309
a_3	0.637	0.216	0.297	0.582

Table 6 Paired comparison matrix among criteria in alternatives

$\tilde{a_1}$	C_1	C_2	C_3	C_4	Weights
C_1	1	3	5	7	0.565
C_2	1/3	1	3	5	0.262
C_3	1/5	1/3	1	3	0.118
C_4	1/7	1/5	1/3	1	0.055
					C.I. = 0.044

$\tilde{a_2}$	C_1	C_2	C_3	C_4	Weights
C_1	1	1/8	3	1	0.111
C_2	8	1	9	9	0.732
C_3	1/3	1/9	1	1/3	0.049
C_4	1	1/9	3	1	0.108
					C.I. = 0.052

$\tilde{a_3}$	C_1	C_2	C_3	C_4	Weights
C_1	1	3	9	7	0.596
C_2	1/3	1	8	3	0.266
C_3	1/9	1/8	1	1/3	0.042
C_4	1/7	1/3	3	1	0.097
					C.I. = 0.042

weights are normalized in the ANP, although this normalization isn't needed in the method proposed.

(Step 3-3) We adjust weights of alternatives in (Step 3-1) with weights of criteria in (Step 3-2). In this example we adjust Table 5 with weights of criteria for each basis $\tilde{a_i}$ in Table 6. Firstly, we adjust Table 5 with weights of criteria for $\tilde{a_1}$ and have Table 7.

Table 7 Adjusted evaluation ratio table about Basis $\tilde{a_1}$

$\tilde{a_1}$	C_1	C_2	C_3	C_4
a_1	0.565	0.262	0.118	0.055
a_2	0.229	1.743	0.036	0.156
a_3	1.393	0.552	0.065	0.294

When we denote Table 5, weights of criteria for $\tilde{a_1}$ in Table 6 and Table 7 with $G = (g_{ij})$, $C_{a_1} = (c_j^{a_1})$ and $G_{a_1} = (g_{ij}^{a_1})$, respectively, it holds

(1) $g_{1j}^{a_1} = c_j^{a_1}$ $(j = 1,2,3,4)$,
(2) $g_{1j}^{a_1} : g_{2j}^{a_1} : g_{3j}^{a_1} = g_{1j} : g_{2j} : g_{3j}$ $(j = 1,2,3,4)$.

Similarly, we adjust Table 5 with Bases $\tilde{a_2}$ and $\tilde{a_3}$ and have Table 8.

When we denote Table 5, weights of criteria for $\tilde{a_2}$ and $\tilde{a_3}$ in Table 6, and tables about a_2 and a_3 in Table 8 with $G = (g_{ij})$, $C_{a_2} = (c_j^{a_2})$, $C_{a_3} = (c_j^{a_3})$, $G_{a_2} = (g_{ij}^{a_2})$ and $G_{a_3} = (g_{ij}^{a_3})$, respectively, it holds

(1) $g_{2j}^{a_2} = c_j^{a_2}$ and $g_{3j}^{a_3} = c_j^{a_3}$ $(j = 1,2,3,4)$,
(2) $g_{1j}^{a_2} : g_{2j}^{a_2} : g_{3j}^{a_2} = g_{1j}^{a_3} : g_{2j}^{a_3} : g_{3j}^{a_3} = g_{1j} : g_{2j} : g_{3j}$ $(j = 1,2,3,4)$.

Table 8 Adjusted evaluation ratio table about Basis $\tilde{a_2}$ and $\tilde{a_3}$

$\tilde{a_2}$	C_1	C_2	C_3	C_4
a_1	0.273	0.110	0.162	0.038
a_2	0.111	0.732	0.049	0.108
a_3	0.672	0.232	0.089	0.203

$\tilde{a_3}$	C_1	C_2	C_3	C_4
a_1	0.241	0.126	0.077	0.018
a_2	0.098	0.839	0.023	0.051
a_3	0.596	0.266	0.042	0.097

(Step 3-4) We synthesize three tables in Tables 7 and 8 and have Table 9, which is the guessed evaluation ratio table of alternatives with respect to criteria under the goal. In order to do this, we calculate geometric means of corresponding components in tables. The validity to adopt geometric means is explained in Section 4.

Table 9 Guessed evaluation ratio table of alternatives with respect to criteria under the goal

Goal	C_1	C_2	C_3	C_4
a_1	0.334	0.154	0.114	0.034
a_2	0.136	1.023	0.034	0.095
a_3	0.823	0.324	0.062	0.180

(Step 4) We combine weights of alternatives in (Step 3-4), i.e., Table 9, with weights of criteria in (Step 2), i.e., Table 3 and have

$$\begin{pmatrix} \text{Weight of } a_1 \\ \text{Weight of } a_2 \\ \text{Weight of } a_3 \end{pmatrix} = \begin{pmatrix} 0.334 & 0.154 & 0.114 & 0.034 \\ 0.136 & 1.023 & 0.034 & 0.095 \\ 0.823 & 0.324 & 0.062 & 0.180 \end{pmatrix} \begin{pmatrix} 0.657 \\ 0.203 \\ 0.094 \\ 0.046 \end{pmatrix} = \begin{pmatrix} 0.263 \\ 0.304 \\ 0.621 \end{pmatrix}.$$

Finally, by normalization, which means the sum of weights is 1, we have Weight of a_1 : Weight of a_2 : Weight of a_3 = 0.221 : 0.256 : 0.523 and a_1 (0.221) $\prec$ a_2 (0.256) $\prec$ a_3 (0.523). The decision maker chooses alternative a_3.

(Step 5) We recommend checking validity of the ratio of weights of alternative obtained in (Step 4). We explain this in Section 5.

4 Validity of the Method Proposed

In this section we deal with a three-level hierarchy with four criteria and three alternatives as in Sections 2 and 3 (see Fig 1). Firstly, we show the validity in the case that alternatives have concrete evaluation values with respect to each criterion of Table 10. This means that 'if we have concrete evaluation value for each alternative with respect to criteria, the final ratio is the same as one with weighted summation method' as mentioned in Section 1.

Firstly, in (Step 2) we obtain a continued ratio of weights $w_{C_1} : w_{C_2} : w_{C_3} : w_{C_4}$ for criteria C_j $(j = 1,2,3,4)$ with a certain technique, which satisfies the condition

$$w_{C_1} + w_{C_2} + w_{C_3} + w_{C_4} = 1.$$

Table 10 Concrete evaluation values of alternatives with respect to criteria

	C_1	C_2	C_3	C_4
a_1	w_{11}	w_{12}	w_{13}	w_{14}
a_2	w_{21}	w_{22}	w_{23}	w_{24}
a_3	w_{31}	w_{32}	w_{33}	w_{34}

This condition is usual in multi-criteria decision-making methods. Next, we calculate a ratio of weights of alternatives with respect to criteria in (Step 3-1). This is represented as Figure 2. Then we obtain Table 11 with constants α_j $(j = 1,2,3,4)$ with paired comparisons according to Table 10. If we use the AHP/ANP, we have $\alpha_j = 1/(w_{1j} + w_{2j} + w_{3j})$ $(j = 1,2,3,4)$, which means normalization.

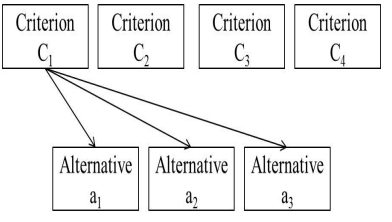


Fig. 2 Evaluation of alternatives for a criterion

Table 11 The ratio of evaluation of alternatives for each criterion

	C_1	C_2	C_3	C_4
a_1	$\alpha_1 w_{11}$	$\alpha_2 w_{12}$	$\alpha_3 w_{13}$	$\alpha_4 w_{14}$
a_2	$\alpha_1 w_{21}$	$\alpha_2 w_{22}$	$\alpha_3 w_{23}$	$\alpha_4 w_{24}$
a_3	$\alpha_1 w_{31}$	$\alpha_2 w_{32}$	$\alpha_3 w_{33}$	$\alpha_4 w_{34}$

We move to (Step 3-2) and calculate weights of criteria in each alternative. This is represented as Figure 3 (cf. Figure 2). The purpose of this step is guessing that $w_{i1} : w_{i2} : w_{i3} : w_{i4}$ for each i $(i = 1,2,3)$ in Table 10 (see Figure 4). Then we have the ratio of evaluation values of criteria occupied in each basis of Table 12 with constants β_i $(i = 1,2,3)$. If we use the ANP, we have $\beta_i = 1/(w_{i1} + w_{i2} + w_{i3} + w_{i4})$ $(i = 1,2,3)$, which means normalization.

Table 12 The ratio of evaluation values of criteria occupied in alternatives

	$\tilde{a_1}$	$\tilde{a_2}$	$\tilde{a_3}$
C_1	$\beta_1 w_{11}$	$\beta_2 w_{21}$	$\beta_3 w_{31}$
C_2	$\beta_1 w_{12}$	$\beta_2 w_{22}$	$\beta_3 w_{32}$
C_3	$\beta_1 w_{13}$	$\beta_2 w_{23}$	$\beta_3 w_{33}$
C_4	$\beta_1 w_{14}$	$\beta_2 w_{24}$	$\beta_3 w_{34}$

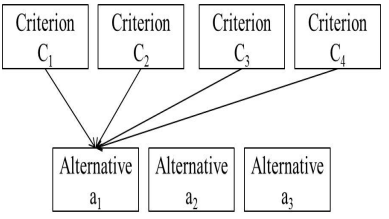


Fig. 3 Evaluation of criteria occupied in $\tilde{a_1}$

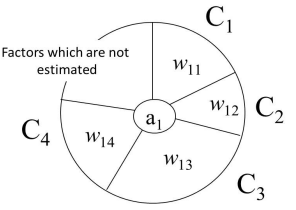


Fig. 4 Evaluation of criteria occupied in $\tilde{a_1}$

Table 13 Adjusted evaluation ratio table about Basis $\tilde{a_k}$ $(k = 1,2,3)$

$\tilde{a_k}$	C_1	C_2	C_3	C_4
a_1	$\beta_k w_{11}$	$\beta_k w_{12}$	$\beta_k w_{13}$	$\beta_k w_{14}$
a_2	$\beta_k w_{21}$	$\beta_k w_{22}$	$\beta_k w_{23}$	$\beta_k w_{24}$
a_3	$\beta_k w_{31}$	$\beta_k w_{32}$	$\beta_k w_{33}$	$\beta_k w_{34}$

In (Step 3-3) we adjust Table 11 with the column for each basis $\tilde{a_i}$ in Table 12 and obtain Table 13.

When we denote Table 11, Table 12 and Table 13 with $k = 1,2,3$ with $G = (g_{ij})$, $C = (c_{ij})$ and $G_{a_k} = (g_{ij}^{a_k})$, respectively, it holds for k $(k = 1,2,3)$

(1) $g_{kj}^{a_k} = c_{jk} = \beta_k w_{kj}$ $(j = 1,2,3,4)$,
(2) $g_{1j}^{a_k} : g_{2j}^{a_k} : g_{3j}^{a_k} = g_{1j} : g_{2j} : g_{3j} = w_{1j} : w_{2j} : w_{3j}$ $(j = 1,2,3,4)$.

We have the same number of tables as alternatives in Table 13 and synthesize those tables in (Step 3-4). In order to do this, we use geometric means method for corresponding components of those tables and have Table 14.

Table 14 Guessed weights of alternatives with respect to criteria under the goal

	C_1	C_2	C_3	C_4
a_1	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{11}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{12}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{13}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{14}$
a_2	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{21}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{22}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{23}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{24}$
a_3	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{31}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{32}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{33}$	$(\Pi_{k=1}^3 \beta_k)^{1/3} w_{34}$

Finally, in (Step 4) we combine weights of alternatives in Table 14 with weights of criteria acquired in (Step 2), which are w_{C_j} $(j = 1,2,3,4)$. When we denote weights of alternatives a_i $(i = 1,2,3)$ with respect to the goal before and after normalizing with W_{a_i} and N_{a_i}, respectively, we have the following:

$$\begin{pmatrix} W_{a_1} \\ W_{a_2} \\ W_{a_3} \end{pmatrix} = (\Pi_{k=1}^3 \beta_k)^{1/3} \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \end{pmatrix} \begin{pmatrix} w_{C_1} \\ w_{C_2} \\ w_{C_3} \\ w_{C_4} \end{pmatrix}$$

and

$$\begin{pmatrix} N_{\mathrm{a}_1} \\ N_{\mathrm{a}_2} \\ N_{\mathrm{a}_3} \end{pmatrix} = \frac{1}{W_{\mathrm{a}_1} + W_{\mathrm{a}_2} + W_{\mathrm{a}_3}} \begin{pmatrix} W_{\mathrm{a}_1} \\ W_{\mathrm{a}_2} \\ W_{\mathrm{a}_3} \end{pmatrix}.$$

Consequently, we have the final answer of the continued ratio $N_{\mathrm{a}_1} : N_{\mathrm{a}_2} : N_{\mathrm{a}_3}$ for weights of alternatives, which is the same as one by weighted summation method with weights of criteria w_{C_j} $(j = 1,2,3,4)$ and Table 10. In fact, it holds

$$N_{\mathrm{a}_1} : N_{\mathrm{a}_2} : N_{\mathrm{a}_3} = W_{\mathrm{a}_1} : W_{\mathrm{a}_2} : W_{\mathrm{a}_3} = \Sigma_{j=1}^4 W_{\mathrm{C}_j} w_{1j} \ : \ \Sigma_{j=1}^4 W_{\mathrm{C}_j} w_{2j} \ : \ \Sigma_{j=1}^4 W_{\mathrm{C}_j} w_{3j}.$$

Now we shall recall the purpose of the method in Section 2 again. We showed the validity of the method in the case that we have a concrete evaluation value table of alternatives with respect to criteria at the starting point. We propose a test of whether we can assume such a table in the next section.

Next, we explain more general case. By this way we can directly calculate the final answer. (Step 1) We have a three-level hierarchy with one goal G, criteria C_j $(j = 1,2,\ldots,n)$ and alternatives a_i $(i = 1,2,\ldots,m)$. (Step 2) We calculate w_{C_j} as weight of a criterion C_j, where $\Sigma_{j=1}^n w_{C_j} = 1$, and set $C = (w_{C_j})$. (Step 3-1) we obtain $m \times n$ matrix $V = (v_{ij})$ and (Step 3-2) we have $n \times m$ matrix $H = (h_{ij})$ with paired comparisons. (Step 3-3) We adjust those matrices and have $H_k = \left(v_{ij} \times h_{jk}/v_{kj}\right)$ $(k = 1,2,\ldots,m)$. (Step 3-4) We synthesize H_k $(k = 1,2,\ldots,m)$ using geometric means method for corresponding components of H_k and obtain

$$W = \left(\left(\Pi_{k=1}^m v_{ij} \times h_{jk}/v_{kj} \right)^{1/m} \right).$$

(Step 4) We normalize the vector WC and have $N = (N_{\mathrm{a}_i})$ with $\Sigma_{i=1}^m N_{\mathrm{a}_i} = 1$. Consequently, N_{a_i} $(i = 1,2,\ldots,m)$ are weights of alternative a_i with respect to G as the final answer. We note that if we obtain a concrete evaluation value ζ of alternative a_i with respect to a criterion C_j for some i and j later, we can reconstruct Table 10 using $w_{ij} = \zeta$.

We have the following theorem which shows validity of the method in the case without a concrete evaluation value table. This is a reason to adopt geometric means method for each component to synthesize tables in (Step 3-4).

Theorem 1. *Let $W = (w_{ij})$ and $W' = (w'_{ij})$ be $m \times n$ matrices with $w_{ij} > 0$ and $w'_{ij} > 0$ for some integers m and n $(m,\ n \geq 2)$. Then the continued ratio $\Sigma_{j=1}^n C_j w_{1j} : \Sigma_{j=1}^n C_j w_{2j} : \cdots : \Sigma_{j=1}^n C_j w_{mj}$ is equal to $\Sigma_{j=1}^n C_j w'_{1j} : \Sigma_{j=1}^n C_j w'_{2j} : \cdots : \Sigma_{j=1}^n C_j w'_{mj}$ for any set $\{C_j\ (j = 1,2,\ldots,n) \mid C_j > 0$ and $\Sigma_{j=1}^n C_j = 1\}$ if and only if there exists a positive number s such that $w'_{ij} = s w_{ij}$ for all i and j $(i = 1,2,\ldots,m,\ j = 1,2,\ldots,n)$.*

Proof. It is sufficient only to prove the necessity condition. We set $T = \Sigma_{i=1}^m \Sigma_{j=1}^n C_j w_{ij}$ and $T' = \Sigma_{i=1}^m \Sigma_{j=1}^n C_j w'_{ij}$. It follows from the assumption that there is a number t such that $(1/T')\Sigma_{j=1}^n C_j w'_{ij} = t(1/T)\Sigma_{j=1}^n C_j w_{ij}$ for all i $(i = 1,2,\ldots,m)$. Then since $\Sigma_{i=1}^m((1/T')\Sigma_{j=1}^n C_j w'_{ij}) = \Sigma_{i=1}^m(t(1/T)\Sigma_{j=1}^n C_j w_{ij})$ and

definitions of T and T', we have $t = 1$. So we have $(1/T')\Sigma_{j=1}^{n} C_j w'_{ij} = (1/T)\Sigma_{j=1}^{n} C_j w_{ij}$ for all i $(i = 1,2,\ldots,m)$.

In particular, we have $(1/T')\Sigma_{j=1}^{n} w'_{ij} = (1/T)\Sigma_{j=1}^{n} w_{ij}$ $(i = 1,2,\ldots,m)$ when $C_1 = \cdots = C_n = 1/n$, and $(1/T')(2(n-1)w'_{ik} + \Sigma_{j\neq k} w'_{ij}) = (1/T)(2(n-1)w_{ik} + \Sigma_{j\neq k} w_{ij})$ when $C_k = 2/3$ and $C_j = 1/(3(n-1))$ $(j \neq k)$. It follows from these two equations that $(1/T')(2n-3)w'_{ik} = (1/T)(2n-3)w_{ik}$ and $w'_{ik} = (T'/T)w_{ik}$ for all i and k $(i = 1,2,\ldots,m,\ k = 1,2,\ldots,n)$ and the theorem was proved. □

Let m and n be integers greater than or equal to 2 and $M(m,n)^+$ be the set of all $m \times n$ matrices in which all components are positive. We write $W \sim W'$ for W and $W' \in M(m,n)^+$ if there exists a positive number s satisfying $W' = sW$. Then the relation $\sim$ is an equivalence relation in $M(m,n)^+$. We choose $W = (w_{ij})$ with $\Sigma_{i=1}^{m}\Sigma_{j=1}^{n} w_{ij} = 1$ as a representative of each equivalence class. Indeed for any matrix $M = (m_{ij}) \in M(m,n)^+$, we can take the representative $(1/m)M \in \overline{M}$ for $m = \Sigma_{i=1}^{m}\Sigma_{j=1}^{n} m_{ij}$ (> 0). When we deal with the method proposed in this paper, it is sufficient to consider these equivalence classes $Cl(M(m,n)^+)$ from Theorem 1.

Therefore, we only need to define a map f_t of t-tuple of $Cl(M(m,n)^+)$ into $Cl(M(m,n)^+)$ in order to synthesize t evaluation ratio matrices in (Step 3-4). Let E_k $(k = 1,2,\ldots,t)$ be equivalent classes of $Cl(M(m,n)^+)$ and $E_k = \overline{W_k} = \overline{(w_{ij}^k)}$. Then, as was mentioned above, we adopted geometric means method for corresponding entries of matrices W_k as f_t. Namely, for $E_k \in Cl(M(m,n)^+)$ and $W_k = (w_{ij}^k) \in E_k$ $(k = 1,2,\ldots,t)$ and we define

$$f_t(E_1,\ E_2,\ \ldots,\ E_t) = \overline{((\Pi_{k=1}^{t} w_{ij}^k)^{1/t})}.$$

Then we remark that it is independent of the choice of the representative $W_k \in E_k$ for any equivalence class $E_k \in Cl(M(m,n))$. Namely, the map f_t is well-defined for each t. Therefore, it is valid to adopt geometric means method in order to synthesize evaluation value matrices.

Finally, we showed how to mix V in (Step 3-1) and H in (Step 3-2) in Section 3. We adjusted V with each column of H in (Step 3-3). On the other hand, we can formally adjust H with each row of V instead. We chose the former because the scale of a criterion to pairly compare alternatives is more precise than the scale of alternative to pairly compare rates of criteria which occupy the alternative.

5 Validity Test of the Final Answer with the Method Proposed

As shown in Section 4 we can always calculate $W = (w_{ij})$ in (Step 3-4) adjusting $V = (v_{ij})$ in (Step 3-1) with $H = (h_{ij})$ in (Step 3-2), but it is easy to see that we don't always guess W. In this section we propose a test to check this. This test is used in (Step 5).

we calculate $H_k = (v_{ij} \times h_{kj}/v_{kj})$ $(k = 1,2,\ldots,m)$ in (Step 3-3). It is easily seen that if we have a concrete evaluation value table W like Table 10 at the beginning, H_k $(k = 1,2,\ldots,m)$ are equal to each other except for an arbitrary constant factor.

So we think that H_k is almost equal to W and calculate weights of alternatives by combining with $C = (w_{C_j})$, respectively. Then we compare the rankings of the alternatives by their weights and if all the ranking is the same, then we consider that W given by synthesizing H_k $(k = 1, 2, \ldots, m)$ is valid and the final answer by W is so.

Now we check the final answer of the example in Section 3. We have the following for bases $\tilde{a}_1$, $\tilde{a}_2$ and $\tilde{a}_3$ from Tables 7, 8 and 3, respectively:

$$\begin{pmatrix} \text{Weight of } a_1 \\ \text{Weight of } a_2 \\ \text{Weight of } a_3 \end{pmatrix} = \begin{pmatrix} 0.565 & 0.262 & 0.118 & 0.055 \\ 0.229 & 1.743 & 0.036 & 0.156 \\ 1.393 & 0.552 & 0.065 & 0.294 \end{pmatrix} \begin{pmatrix} 0.657 \\ 0.203 \\ 0.094 \\ 0.046 \end{pmatrix} = \begin{pmatrix} 0.438 \\ 0.512 \\ 1.047 \end{pmatrix},$$

$$\begin{pmatrix} \text{Weight of } a_1 \\ \text{Weight of } a_2 \\ \text{Weight of } a_3 \end{pmatrix} = \begin{pmatrix} 0.273 & 0.110 & 0.162 & 0.038 \\ 0.111 & 0.732 & 0.049 & 0.108 \\ 0.672 & 0.232 & 0.089 & 0.203 \end{pmatrix} \begin{pmatrix} 0.657 \\ 0.203 \\ 0.094 \\ 0.046 \end{pmatrix} = \begin{pmatrix} 0.219 \\ 0.231 \\ 0.507 \end{pmatrix},$$

$$\begin{pmatrix} \text{Weight of } a_1 \\ \text{Weight of } a_2 \\ \text{Weight of } a_3 \end{pmatrix} = \begin{pmatrix} 0.241 & 0.126 & 0.077 & 0.018 \\ 0.098 & 0.839 & 0.023 & 0.051 \\ 0.596 & 0.266 & 0.042 & 0.097 \end{pmatrix} \begin{pmatrix} 0.657 \\ 0.203 \\ 0.094 \\ 0.046 \end{pmatrix} = \begin{pmatrix} 0.192 \\ 0.239 \\ 0.454 \end{pmatrix}$$

and we have Weight of $a_1 \prec$ Weight of $a_2 \prec$ Weight of a_3 for all bases. As the result, we accept the final answer a_1 (0.221) $\prec a_2$ (0.256) $\prec a_3$ (0.523).

6 Conclusions

In this paper we proposed a decision-making method, which calculates the ratio of evaluation value of alternatives with respect to a goal with paired comparisons. It is natural to use paired comparisons in this method because we are interested in ratio of weights of items, which are alternatives or criteria. We used paired comparisons in the AHP/ANP in this paper, but it is a problem in the future to use what kind of measure with paired comparisons in this method in order to achieve decision-makers' purpose. Moreover, we showed a validity test for the final answer in Section 5, but we need to research more appropriate test. Finally, we note that we can apply this method for a hierarchy with sub-criteria, though we dealt with only three-level hierarchy of the goal, criteria and alternatives in this paper.

Acknowledgements. The author would like to express his sincere gratitude to Professor Thomas L. Saaty, Professor Rozann W. Saaty and Professor Luis Vargas at University of Pittsburgh for their helpful and useful advice in the AHP/ANP. He also thanks Professor Eizo Kinoshita at Meijo University and Professor Kazuyuki Sekitani at Shizuoka University for their useful advice and considerable encouragement. This work was supported by Tokyo University of Science, Suwa.

References

1. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
2. Saaty, T.L.: The Analytic Network Process. RWS Publications, Pittsburgh (1996)

A Stackelberg Location on a Network with Fuzzy Random Demand Quantities Using Possibility Measure

Takeshi Uno, Hideki Katagiri, and Kosuke Kato

Abstract. This paper focuses on a competitive facility location problem between leader and follower on a network with demands whose weights are given uncertainly and vaguely. By representing them as fuzzy random variables, the optimal location problem can be formulated as a fuzzy random programming problem for finding Stackelberg equilibrium. For solving the problem, it is reformulated as the problem to find the optimal solutions maximizing a degree of possibility under some chance constraint for the leader. Theorems for its complexity are shown based upon the characteristics of the facility location.

1 Introduction

Competitive facility location problem (CFLP) is one of optimal location problems for commercial facilities, e.g. shops and stores, and an objective of most CFLPs is to obtain as many buying powers (BPs) from customers as possible. Mathematical studies on the CFLPs were originated by Hotelling [7]. He considered the CFLP under the conditions that (i) customers are uniformly distributed on a line segment, (ii) each of decision makers (DMs) will locate her/his own facility on the line segment

Takeshi Uno
Institute of Socio-Arts and Sciences, The University of Tokushima, 1-1,
Minamijosanjima-cho, Tokushima-shi, Tokushima, 770-8502, Japan
e-mail: uno@ias.tokushima-u.ac.jp

Hideki Katagiri
Graduate School of Engineering, Hiroshima University, 1-4-1,
Kagamiyama, Higashihiroshima-shi, Hiroshima, 739-8527, Japan
e-mail: katagiri-h@hiroshima-u.ac.jp

Kosuke Kato
Faculty of Applied Information Science, Hiroshima Institute of Technology, 2-1-1,
Miyake, Saeki-ku, Hiroshima, 731-5193, Japan
e-mail: k.katoh.me@it-hiroshima.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 67–75.
springerlink.com

that there are no facilities, and (iii) all customers only use the nearest facility. Then, his CFLP can be represented as an optimal location problem for finding Nash equilibrium, called Nash location problem (NLP). As an extension of Hotelling's NLP, Wendell and McKelvey [25] assumed that there exist customers on a finite number of points, called demand points (DPs), and considered an NLP on a tree network whose vertices are DPs.

On the other hand, based upon the above NLP on a network, Hakimi [5] considered the CFLP with two types of DMs; the leader, who first locates her/his facilities, and the follower, who next locates her/his facilities. Then, his CFLP for the leader can be represented as an optimal location problem for finding Stackelberg equilibrium, called Stackelberg location problem (SLP). For details of Hakimi's SLP and their applications, the readers can refer to the book of Miller et al. [13]. As an extension of Hakimi's SLP, SLPs on a plane are considered by Drezner [3], Uno et al. [19, 20], etc. Another type of SLP based on maximal covering is considered by Plastria and Vanhaverbeke [15].

In the above studies of CFLPs, the demands of customers for facilities are represented as definite values. We consider some uncertainty and vagueness with demand for facilities. For the uncertainty, facility location model with random demand in a noncompetitive environment is considered by Wagnera et al. [23]; for the details of location models with random demands, the reader can refer to the study of Berman and Krass [2]. For CFLPs with random demands, Shiode and Drezner [17] considered an SLP on a tree network, and Uno et al. [19] considered a CFLP on a plane. On the other hand, for the vagueness, facility location model with fuzziness in a noncompetitive environment is considered by Moreno Pérez et al. [14], which represented the demands as fuzzy numbers proposed by Dubois and Prade [4]. Recently, the decision-making problems in environments including both uncertainty and vagueness are studied. Kwakernaak [12] proposed the fuzzy random variable representing both fuzziness and randomness. For the details of fuzzy random variable, the reader can refer to the book of Kruse and Meyer [11]. Fuzzy random programming and its distribution problems are considered by Wang and Qiao [24], and Qiao and Wang [16]. For the recent studies of fuzzy random programming problems, Katagiri et al. [10] considered multiobjective fuzzy random linear programming, and Ammar [1] considered fuzzy random multiobjective quadratic programming. Uno et al. [20] considered CFLPs with weights represented as fuzzy random numbers, and Uno et al. [21] considered SLPs with demands on a tree network, whose sites are represented as fuzzy random variables.

In this paper, we consider the SLP proposed by Uno et al. [22], with demands on a tree network whose quantities are represented as fuzzy random numbers. By representing the demands as fuzzy random variables, the SLP can be formulated as a fuzzy random programming problem. For solving it, they introduced the α-level set for their fuzzy random demands. However, it is often difficult for the DM to give a suitable α-level set for the problem, and then their CFLP needs to be introduced another evaluation for the fuzziness of demands. Then, we propose reformulating their SLP as the problems to find their optimal solutions maximizing a degree of possibility under some chance constraint. For details of possibility measure for fuzzy

random variables, the reader can refer to the studies of Katagiri et al. [8, 9]. Then, we can reformulate it to a deterministic programming problem on a tree network, and show its complexity and solution method based upon the characteristics of SLP.

The remaining structure of this article is organized as follows: The next section devotes to introducing the definition of fuzzy random variables. In Section 3, we formulate the SLP on a tree network with demands whose sites are given uncertainly and vaguely as an SLP with fuzzy random distances. For solving it, we reformulate it as the problems to find their optimal solutions maximizing a degree of possibility under some chance constraint in Section 4. In Section 5, we show its complexity and solution method based upon the solution method for the conventional SLP. Finally, conclusions and future studies are summarized in Section 6.

2 Fuzzy Random Variable

Let A be fuzzy number and $\mu_A : \mathbf{R} \to [0,1]$ be membership function of A, where $\mathbf{R}$ is the set of real numbers. For $\alpha \in (0,1]$, the α-level set of A is represented as the following equation:

$$A_\alpha \equiv \{x |\ \mu_A(z) \geq \alpha\} \tag{1}$$

In this paper, we use the following definition of fuzzy random variable, suggested by Kruse and Meyer [11]:

Definition 1. Let (Ω, B, P) be a probability space, where Ω, B, and P are a sample space, σ-algebra, and a probability measure function, respectively. Let $\mathscr{F}(\mathbf{R})$ be the set of fuzzy numbers with compact supports, and Ξ a measurable mapping $\Omega \to \mathscr{F}(\mathbf{R})$. Then Ξ is a fuzzy random variable if and only if given $\omega \in \Omega$, its α-level set $\Xi_\alpha(\omega)$ is a random interval for any $\alpha \in (0,1]$.

Fig. 1 illustrates an example of fuzzy random variables for representing BP per day for weather, whose randomness is represented by weather and whose fuzziness is included in the BP for each case of weather.

3 Formulation of SLP with Fuzzy Random Quantity Demanded

We consider the SLP on a weighted tree $T = (V, E)$, which is a simple graph, where V and E are the sets of vertices and edges, respectively. For each vertex $v \in V$ and edge $e \in E$, we associate weights $w(v), l(e) \geq 0$, respectively, where $w(v)$ means the BP of the weight on v for facilities and $l(e)$ the length of e. Fig. 2 illustrates an example of tree networks of the SLP.

In the tree T, we consider the case that each $w(v)$, $v \in V$ is given as the following fuzzy random variable:

- Its randomness is given by γ scenarios, whose probabilities are denoted by $p_1, p_2, \ldots, p_\gamma > 0$.

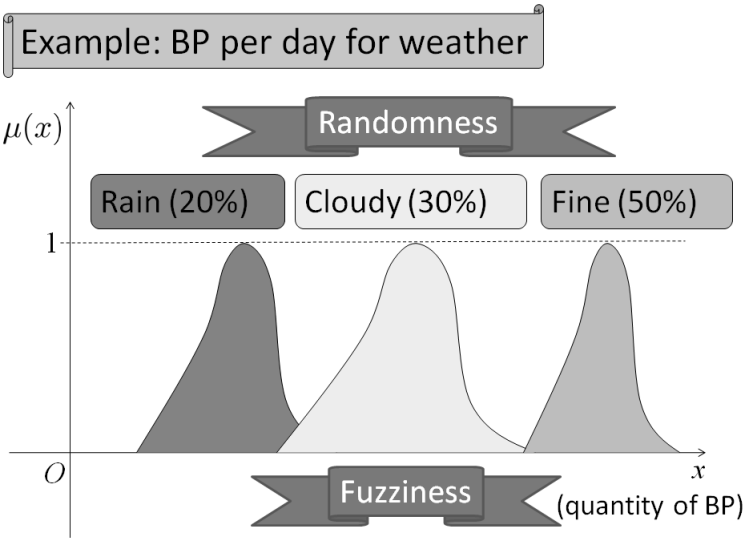


Fig. 1 An Example of Fuzzy Random Variables

- For each scenario $s = 1, 2, \ldots, \gamma$, its fuzziness is given as fuzzy number $w_s(v) \in \mathscr{F}(\mathbf{R})$ whose membership function is denoted by $\mu_{w_s(v)}(z)$, where $\mu_{w_s(v)}(z) = 0$ for any $z \leq 0$.

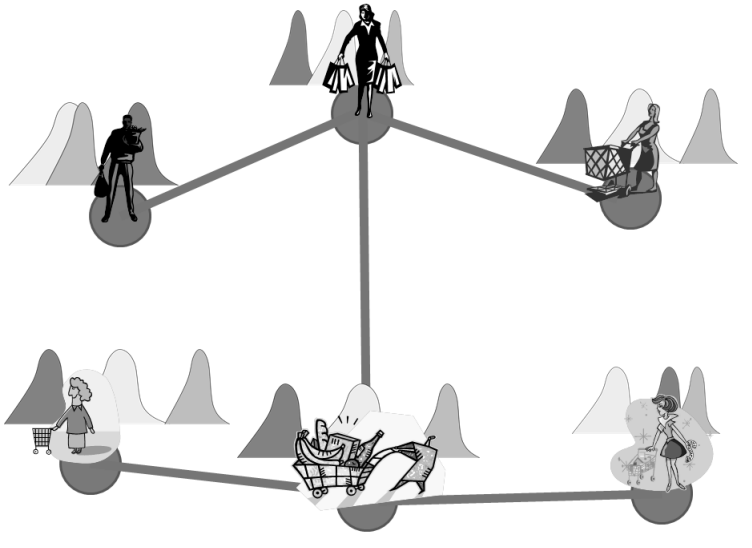


Fig. 2 An Example of Tree Networks of the SLP

An example of fuzzy random weights is shown in Fig. 2.

Let q and r be the given numbers of facilities located by the leader and the follower, respectively. Let $x_1, x_2, \ldots, x_q \in T$ be the sites of the leader's facilities and $X_q = \{x_1, x_2, \ldots, x_q\}$. Similarly, let $y_1, y_2, \ldots, y_r \in T$ be the sites of the follower's facilities and $Y_r = \{y_1, y_2, \ldots, y_r\}$. We assume that each of customers only use the nearest facility, and the facility providing some service to customers on v can obtains $w(v)$ from $v \in V$. If two or more facilities are the same distances to a customer, one of the leader's facilities can obtain her/his BP.

Let $W_{FR}(X_q, Y_r)$ be the sum of obtaining BPs of the leader's facilities from the customers. Note that $W_{FR}(X_q, Y_r)$ is represented as a fuzzy random number. The objective of each DM is defined to maximize her/his obtaining BPs. Since the sum of obtaining BPs of all facilities is constant, the objective of the follower can be represented as minimizing the sum of the leader's obtaining BPs. For given location $X_q \in T^q = T \times \cdots \times T$, the optimal location problem for the follower, called $(X_q|r)$-medianoid problem, can be formulated as follows:

$$\left.\begin{array}{l} \text{minimize } A_f := W_{FR}(X_q, Y_r) \\ \text{subject to } Y_r \in T^r. \end{array}\right\} \tag{2}$$

Let $Y_r^*(X_q)$ be the optimal solution of $(X_q|r)$-medianoid problem. Then, the proposed SLP, called $(q|r)$-centroid problem, can be formulated as follows:

$$\left.\begin{array}{l} \text{maximize } W_{FR}(X_q, Y_r^*(X_q)) \\ \text{subject to } X_q \in T^q. \end{array}\right\} \tag{3}$$

4 Reformulation of CFLP Using Possibility Measure

Taking account of vagueness of human judgment of the DMs, the leader gives her/his fuzzy goal "the leader hopes to make the leader's objective function value y roughly larger than l_1 but certainly larger than l_0" as the following membership function:

$$\mu_{G_l}(y) = \begin{cases} 1, & \text{if } y \geq l_1, \\ \dfrac{l_1 - y}{l_1 - l_0}, & \text{if } l_0 < y < l_1, \\ 0, & \text{if } y \leq l_0. \end{cases} \tag{4}$$

An example of setting l_1 and l_0 is that $l_0 = 0$ and l_1 is an expectation value of the sum of all BPs whose membership value is one. Using the leader's fuzzy goal, we give the follower's fuzzy goal as the following membership function:

$$\mu_{G_f}(y) = \begin{cases} 1, & \text{if } y \leq l_0, \\ \dfrac{y - l_0}{l_1 - l_0}, & \text{if } l_0 < y < l_1, \\ 0, & \text{if } y \geq l_1. \end{cases} \tag{5}$$

Given the follower's membership function of fuzzy random variable A_f and that of a fuzzy goal G_f, a degree of possibility that her/his obtaining BPs satisfy the fuzzy goal can be defined as follows:

$$\prod_{A_f}(G_f) = \sup_y \min\left\{\mu_{A_f}(y), \mu_{G_f}(y)\right\}. \tag{6}$$

In this paper, we use the following chance constraint proposed by Katagiri et al. [8]:

$$Pr\left(\prod_{A_f}(G_f) \geq h\right) \geq \theta_f, \tag{7}$$

where h is an auxiliary variable satisfying $0 \leq h \leq 1$, and $\theta_f \in (0,1)$ is a satisfying probability level, given by the leader. Then, (2) can be reformulated as the following problem:

$$\left.\begin{array}{ll} \text{maximize} & h \\ \text{subject to} & Pr\left(\prod_{A_f}(G_f) \geq h\right) \geq \theta_f, \\ & Y_r \in T^r. \end{array}\right\} \tag{8}$$

Let $Y_r^{(G_f,\theta_f)}(X_q)$ be the optimal solution of (8) for X_q, and we replace the leader's objective function of (3) with the following function:

$$A_l(X_q) := W_{FR}(X_q, Y_r^{(G_f,\theta_f)}(X_q)) \tag{9}$$

Then, a degree of possibility that her/his obtaining BPs satisfy the fuzzy goal can be defined as follows:

$$\prod_{A_l(X_q)}(G_l) = \sup_y \min\left\{\mu_{A_l(X_q)}(y), \mu_{G_l}(y)\right\}. \tag{10}$$

Using the chance constraint (7) with $\theta_l \in (0,1)$, a satisfying probability level given by the leader, (3) can be reformulated as the following problem:

$$\left.\begin{array}{ll} \text{maximize} & h \\ \text{subject to} & Pr\left(\prod_{A_l(X_q)}(G_l) \geq h\right) \geq \theta_l, \\ & X_q \in T^q. \end{array}\right\} \tag{11}$$

5 Complexity of the SLP

For cases that the tree network does not include any fuzzy random weights, (2) and (3) can be reduced to conventional medianoid and centroid problems, respectively, which are NP-hard if $q \geq 2$ proven by Hakimi [6] and Spoerhase and Wirth [18]. Therefore, the following theorems are apparently satisfied:

Lemma 1. *For any $X_q \in T^q$ with $q \geq 2$, (8) is NP-hard.*

Theorem 1. *For any $q \geq 2$, (11) is NP-hard.*

Then, we consider (8) and (11) for the case $q = 1$. We first show the following two lemmas for solving (8).

Lemma 2. *If X_1 is on any vertex $v \in V$, then one of $Y_r^{(}X_1)$ can be given by locating all r facilities on the opposite vertices of the edges adjacent to v for any $\theta_f, \theta_l \in (0,1)$.*

Proof. Note that any tree can be cut to several trees by removing any one non-leaf vertex or edge. A lower DM's facility can obtain all BPs on the tree that is cut at a point between her/his facility and the upper DM's facility and includes her/his facility. The best location of the lower DM is clearly so as not to put any nodes between her/his facilities and v. This means that one of $Y_r^{(G_f,\theta_f)}(X_1)$ can be represented by locating her/his r facilities on the set of the above points.

Similarly to the above proof, the following lemma can be shown.

Lemma 3. *If X_1 is on point z in any edge $e \in E$, then one of $Y_r^{(G_f,\theta_f)}(X_1)$ is to locate facilities on both vertices adjacent to e if $r \geq 2$, or either vertex if $r = 1$, for any $\theta_f, \theta_l \in (0,1)$.*

Next we consider (11) for the case $q = 1$.

Theorem 2. *The optimal solution for (11) with $q = 1$ is to locate it on one of the vertices for any $\theta_f, \theta_l \in (0,1)$.*

Proof. We show the proof of the theorem by the reduction to absurdity. We assume the upper DM locates on any edge $e \in E$. If $r \leq 2$, the lower DM can reduce the objective function value of (11) to zero by locating her/his two facilities at both points adjacent to e. On the other hand, if $r = 1$, the optimal location of the lower DM can be found by Lemma 3 and its candidates are only two points. Whichever is optimal for the lower DM, the upper DM can obtain more BPs by locating at the vertex than that on e. These contradict the optimality of (11).

Note that we can show the proofs of Lemmas 2, 3, and Theorem 2 in a similar way of those of the conventional medianoid and centroid problems.

Finally, we consider the complexity for (11) for the case $q = 1$. From Theorem 2, we can find an optimal solution of (11) by examining all vertices. For the case that the upper DM locates her/his one facility on each vertex $v \in V$, we need to solve (8). From Lemma 2, (8) for each location can be solved by examining all the opposite vertices of the edges adjacent to v. Let $|E|$ denote the number of edges. Then, for all locations of the upper DM, the total number of the examination is $2|E|$. This means that (11) can be solved in polynomial time.

6 Conclusions and Future Studies

In this paper, we have considered a Stackelberg location problem on a network with weights including uncertainly and vaguely. For formulating the Stackelberg location problem with the fuzzy random variables, it is reformulated as the problem to find the optimal solutions maximizing a degree of possibility under some chance constraint for each of the leader and the follower. Its complexity have been shown based upon the characteristics of the Stackelberg location.

This paper shows that (8) and (11) with $q \geq 2$ are NP-hard. To propose an efficient solution method for these problems are an important future study. Moreover, extending the SLP to that of general networks is also an important future study.

References

1. Ammar, E.E.: On fuzzy random multiobjective quadratic programming. European Journal of Operational Research 193, 329–340 (2009)
2. Berman, O., Krass, K.: Facility location with stochastic demands and congestion. In: Drezner, Z., Hamacher, H.W. (eds.) Facility Location: Application and Theory, pp. 329–372. Springer, Berlin (2001)
3. Drezner, Z.: Competitive location strategies for two facilities. Regional Science and Urban Economics 12, 485–493 (1982)
4. Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. Journal Systems Science. 9, 613–626 (1978)
5. Hakimi, S.L.: On locating new facilities in a competitive environment. European Journal of Operational Research 12, 29–35 (1983)
6. Hakimi, S.L.: Locations with spatial interactions: Competitive locations and games. In: Mirchandani, P.B., Francis, R.L. (eds.) Discrete Location Theory. Discrete Mathematics and Optimization, pp. 439–478. Wiley Interscience (1990)
7. Hotelling, H.: Stability in competition. The Economic Journal 30, 41–57 (1929)
8. Katagiri, H., Sakawa, M., Hiroaki, I.: Fuzzy random bottleneck spanning tree problems using possibility and necessity measures. European Journal of Operational Research 152, 88–95 (2004)
9. Katagiri, H., Sakawa, M., Kato, K., Nishizaki, I.: A fuzzy random multiobjective 0-1 programming based on the expectation optimization model using possibility and necessity measures. Mathematical and Computer Modelling 40, 411–421 (2004)
10. Katagiri, H., Sakawa, M., Kato, K., Nishizaki, I.: Interactive multiobjective fuzzy random linear programming: Maximization of possibility and probability. European Journal of Operational Research 188, 530–539 (2008)
11. Kruse, R., Meyer, K.D.: Statistics with vague data. D. Riedel Publishing Company, Dordrecht (1987)
12. Kwakernaak, H.: Fuzzy random variables-I. definitions and theorems. Information Sciences 15, 1–29 (1978)
13. Miller, T.C., Friesz, T.L., Tobin, R.L.: Equilibrium facility location on networks. Springer, Berlin (1996)
14. Moreno Pérez, J.A., Marcos Moreno Vega, L., Verdegay, J.L.: Fuzzy location problems on networks. Fuzzy Sets and Systems 142, 393–405 (2004)
15. Plastria, F., Vanhaverbeke, L.: Discrete models for competitive location with foresight. Computers & Operations Research 35, 683–700 (2008)

16. Qiao, Z., Wang, G.: On solutions and distribution problems of the linear programming with fuzzy random variable coefficients. Fuzzy Sets and Systems 58, 120–155 (1993)
17. Shiode, S., Drezner, Z.: A competitive facility location problem on a tree network with stochastic weights. European Journal of Operational Research 149, 47–52 (2003)
18. Spoerhase, J., Wirth, H.C.: (r, p)-centroid problems on paths and trees. Theoretical Computer Science 410, 5128–5137 (2003)
19. Uno, T., Katagiri, H., Kato, K.: Competitive facility location with random demands. IAENG Transactions on Engineering Technologies 3, 83–93 (2009)
20. Uno, T., Katagiri, H., Kato, K.: Competitive facility location with fuzzy random demands. IAENG Transactions on Engineering Technologies 5, 99–108 (2010)
21. Uno, T., Katagiri, H., Kato, K.: A Stackelberg Location Problem on a Tree Network with Fuzzy Random Demands. In: Phillips-Wren, G., Jain, L.C., Nakamatsu, K., Howlett, R.J. (eds.) IDT 2010. SIST, vol. 4, pp. 581–588. Springer, Heidelberg (2010)
22. Uno, T., Katagiri, H., Kato, K.: A competitive facility location problem on a network with fuzzy random weights. Engineering Letters 19, 143–146 (2011)
23. Wagnera, M.R., Bhaduryb, J., Penga, S.: Risk management in uncapacitated facility location models with random demands. Computers & Operations Research 36, 1002–1011 (2009)
24. Wang, G.Y., Qiao, Z.: Linear programming with fuzzy random variable coefficients. Fuzzy Sets and Systems 57, 295–311 (1993)
25. Wendell, R.E., McKelvey, R.D.: New perspective in competitive location theory. European Journal of Operational Research 6, 174–182 (1981)

A Structural Analysis Based on Similarity between Fuzzy Clusters and Its Application to Evaluation Data

Ryunosuke Chiba, Toshiaki Furutani, and Mika Sato-Ilic

Abstract. This paper presents a similarity of fuzzy clusters based analysis for consumer preference evaluation data. We apply this method to multiple types of evaluation data. The merit of this method is to capture the latent structure of consumer preferences represented by fuzzy clusters and the relation between these fuzzy clusters as similarity. In addition, due to the presence of identical fuzzy clusters over the different industries, we can compare the difference between the industries through the same subject evaluation by using the scale of the fuzzy clusters. We show a better performance from the use of our proposed method with several numerical examples.

1 Introduction

The internet consists of various types of information the amount of which is increasing every year. This large amount and variety of data has interested various industrial companies as well as fed the ever expanding need of the public for

Ryunosuke Chiba
Graduate School of Systems and Information Engineering, University of Tsukuba,
Tsukuba, Ibaraki 305-8573
e-mail: s1120540@sk.tsukuba.ac.jp

Toshiaki Furutani
Graduate School of Systems and Information Engineering, University of Tsukuba,
Tsukuba, Ibaraki 305-8573
e-mail: r1020674@risk.tsukuba.ac.jp

Mika Sato-Ilic
Faculty of Engineering, Information and Systems, University of Tsukuba,
Tsukuba, Ibaraki 305-8573
e-mail: mika@risk.tsukuba.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 77–87.
springerlink.com

information exchange. Internet sites are not evaluated by experts but instead the target audience and the public become the evaluator of the online products or stores [1], [2] also recently the data mode or accuracy of estimation of this public opinion observed through web-based surveys have been discussed [5], [6], [10].

The public can actually visited the targeted stores and evaluate based upon their firsthand opinions. Since the public rather than experts are making the assessment, the evaluations fit one's preferences and there are also evaluators with contrary views. In order to use the information from such preference assessment of sites effectively, it is important to capture the latent features of the information, categorize subjects based on their preferences, and identify the obtained latent features to the categorized clusters [9]. Therefore, the purpose of this study is to propose a method to capture this latent structure of the evaluation data as fuzzy clusters [4], [7], [8], and through the fuzzy clusters to identify the features of the various categorized subjects. In addition, using the same scales of degree of belongingness of subjects to fuzzy clusters, temporal difference over the different industries are captured through the similarity of fuzzy clusters.

This paper consists of the following sections: In section 2, we describe a cluster difference scaling method with a fuzzy approximation strategy [3]. In section 3, we propose a method to capture the similarity of fuzzy clusters that obtains the temporal difference over the several industries. In section 4, we show several simulation results and section 5 shows numerical examples of our proposed method. In section 6, we state several conclusions.

2 A Cluster Difference Scaling Method with a Fuzzy Approximation Strategy Introduction

The cluster difference scaling method [3] is shown as follows:

$$\sigma_{kl}^2(E,X) = \sum_{i \in J_k} \sum_{j \in J_l} w_{ij} \left(\delta_{ij} - d_{kl}(X)\right)^2. \tag{1}$$

where E is a $n \times K$ matrix where n shows a number of objects and K is a number of clusters. E shows an indicator matrix in which each element e_{ik} is 0 or 1, if an object i belongs to a cluster k, then e_{ik} is 1, otherwise it is 0. X is a $K \times p$ matrix whose rows show vectors of cluster centers with respect to p variables. w_{ij} shows the weight between objects i and j. δ_{ij} shows dissimilarity between objects i and j which are in clusters k and l respectively. $d_{kl}(X)$ shows a Euclidean distance between kth and lth lows in a matrix X. The purpose of this method is to obtain an optimal partition matrix E and cluster centers X which minimizes equation (1). As an extension of equation (1), the following equation is derived by changing from the hard partition matrix E to a fuzzy partition matrix U where each element u_{ik} has a value in a range from 0 to 1.

$$\sigma_{kl}^2(U,X) = \sum_{i=1}^{n}\sum_{j=1}^{n} u_{ik}^q u_{jl}^q \omega_{ij} \left(\delta_{ij} - d_{kl}(X)\right)^2, \tag{2}$$

where u_{ik} satisfies the following conditions:

$$\sum_{k=1}^{K} u_{ik} = 1, \ ^\forall i, \quad u_{ik} \in [0,1], \ ^\forall i,k. \tag{3}$$

$q(q > 1)$ is a parameter to control degree of fuzziness of clustering and it is given in advance.

3 A Structural Analysis Based on Similarity between Fuzzy Clusters

We define a similarity between a pair of fuzzy clusters k and l as follows:

$$\omega_{kl} \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ikjl}, \tag{4}$$

where

$$g_{ikjl} \equiv u_{ik}^q u_{jl}^q s_{ij}, \tag{5}$$

where s_{ij} shows similarity between objects i and j. Equation (4) has a similar idea to represent a similarity of fuzzy clusters shown in equation (2), that is, in equation (4), we obtain ω_{kl} which is a similarity between fuzzy clusters k and l over the objects rescaled through degree of belongingness which is represented as g_{ikjl} shown in equation (5). However, there is a difference in the way we use the similarity between objects i and j replaced by the weight shown in equation (2). Therefore, we can directly reduce the similarity structure of objects i and j to a similarity of fuzzy clusters k and l through the degree of belongingness in equation (4). This has the merit of using our definition shown in equation (4). Our target data is 3-way temporal data which consists of objects, variables, and situations. The data is represented as follows:

$$X = \left(X^{(1)}, \cdots X^{(T)}\right)', \ X^{(t)} = \left(x_{ia}^{(t)}\right), \ i = 1, \cdots, n, \ a = 1, \cdots, p, \ t = 1, \cdots, T. \tag{6}$$

Where T shows a number of situations. We calculate distance between objects from the data shown in equation (6) as follows:

$$D = \left(d_{ij}\right), \ i,j = 1, \cdots nT, \tag{7}$$

where d_{ij} is a distance between ith and jth rows in X shown in equation (6). Applying a fuzzy clustering method to the distance matrix shown in equation (7), we obtain the fuzzy clustering result as follows:

$$U = (U^{(1)}, \cdots, U^{(T)})', \ U^{(t)} = \left(u_{ik}^{(t)}\right), \ i = 1, \cdots, n, \ k = 1, \cdots, K, \ t = 1, \ldots, T. \quad (8)$$

From objects of $X^{(t)} = \left(x_{ia}^{(t)}\right)$, we obtain a distance matrix as $D^{(t)} = \left(d_{ij}^{(t)}\right)$. The following monotone transformation is used for the transformation from dissimilarity to similarity at t -th situation.

$$s_{ij}^{(t)} = 1 - \left(\frac{d_{ij}^{(t)}}{\max_{i,j}\left(d_{ij}^{(t)}\right)} \right), \ i, j = 1, \cdots, n, \ t = 1, \cdots, T. \quad (9)$$

Then using equations (8) and (9), we define the similarity between fuzzy clusters k and l at t-th situation as follows:

$$\omega_{kl}^{(t)} = \sum_{i=1}^{n}\sum_{j=1}^{n} u_{ik}^{(t)} u_{jl}^{(t)} s_{ij}^{(t)}, \quad k,l=1, \cdots, K, \ t = 1, \cdots, T. \quad (10)$$

$\omega_{kl}^{(t)}$ denotes the similarity between clusters k and l at t-th situation. $u_{ik}^{(t)}$ is degree of belongingness of an object i to a cluster k at t-th situation which satisfies the following conditions.

$$\sum_{k=1}^{K} u_{ik}^{(t)} = 1, \ ^{\forall}i, \quad u_{ik}^{(t)} \in [0,1], \ ^{\forall}i, k, \ t = 1, \cdots, T. \quad (11)$$

4 Simulation Study

We generated artificial data for 1000 objects with respect to bivariate and each value is generated as a normal random number. Each first half object has the same variance and the variances of the later half objects are changed so as to assume that the overlapping statuses of two clusters in this data have changed. Figures 1-7 show data distribution with different variances of the cluster. For the data shown in figures 1-7, we calculate the distance shown in equation (7) when $t = 1$ and apply a fuzzy clustering method called FANNY [4] to this distance and obtained a fuzzy clustering result shown in equation (8) when $t = 1$. The results of similarity between two fuzzy clusters shown in equation (10) when $t = 1$ are shown in table 1. From this table, it can be seen that according to increase in the value of variance, figures 1-7, and shows of the cluster, the similarity of the two clusters also increase.

This is an adaptable result from the data distributions shown in the validity of our definition of similarity between fuzzy clusters.

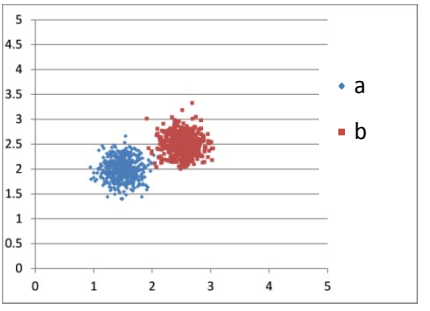


Fig. 1 Variance of group b is 0.2

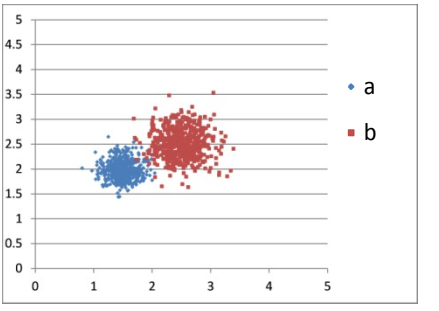


Fig. 2 Variance of group b is 0.3

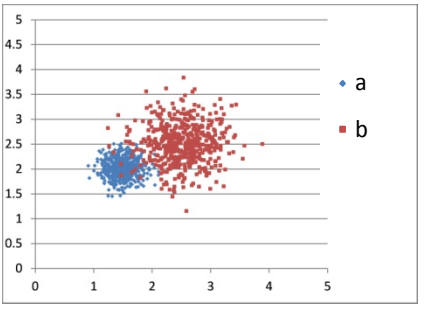


Fig. 3 Variance of group b is 0.4

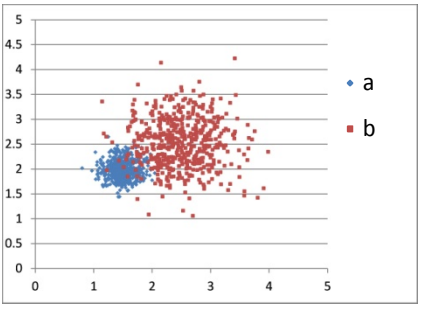


Fig. 4 Variance of group b is 0.5

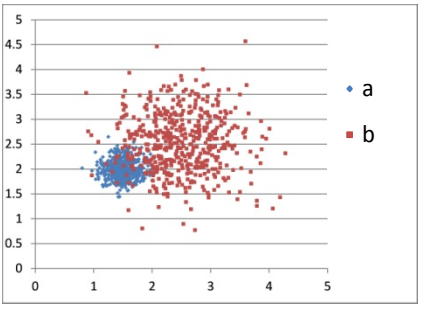


Fig. 5 Variance of group b is 0.6

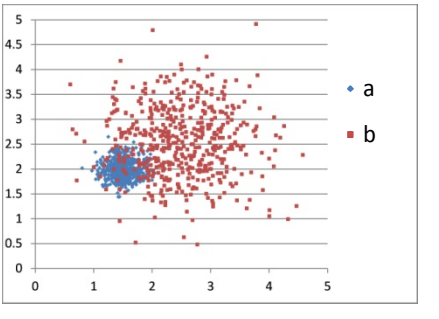


Fig. 6 Variance of group b is 0.7

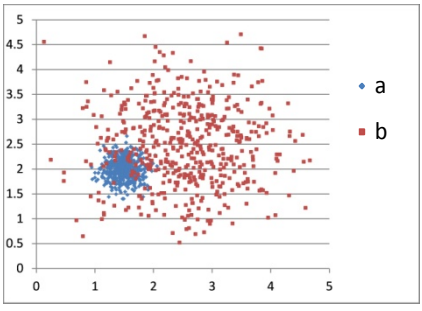


Fig. 7 Variance of group b is 0.8

Table 1 Results of the simulations

Variance of b	ω_{12} $(\times 10^8)$	Mean of ω $(\times 10^8)$	Standard Deviation of ω $(\times 10^7)$
0.2	14.8	17.8	25.8
0.3	15.9	18.1	33.2
0.4	16.2	18.2	53.9
0.5	16.7	18.8	73.2
0.6	17.3	19.5	51.7
0.7	17.6	20.1	59.3
0.8	17.8	20.3	62.6

5 Numerical Examples

We use evaluation data which consists of 46 subjects (34 males, 12 females) with respect to five variables "taste", "cost", "frequency", "service", "quantity" over 4 kinds of foods such as "hamburgers", "rice", "noodles", and "doughnuts ". These 4 kinds of foods represent our 4 types of industries. The subjects were asked if they visited a restaurant for each these types of food, how much they would consider these variables. The degree of the consideration is represented as scores for the four variables "taste", "cost", "service", and "quantity" which are written by the subjects with values from 0 to 5, where 0 is the lowest and 5 is the highest rating. For the variable "frequency", 0 means the subject does not visit the restaurants at all and 5 means that the subject visits the restaurants quite often. The ages of the subjects are from 19 to 29 and they are university students. The subjects were asked to evaluate about the 4 fast-food industries with respect to the five variables. External information for the subjects was obtained on gender, age, and membership in clubs such as sports clubs or cultural interest clubs. The data are shown in tables 2 and 3.

Using the data shown in Table 3, we calculated the distance shown in equation (7). Next, applying a fuzzy clustering method named FANNY [4] to this distance matrix we obtained a fuzzy clustering result shown in equation (8). The number of clusters is assumed to be 4 and table 4 shows the result of the fuzzy clustering.

From this table, it can be seen that we can compare the results among 4 industries due to the same four fuzzy clusters. In other words, since we fix the same clusters, using this measure of degree of belongingness of the subjects to clusters, we can compare the latent structures of data for 4 industries. Tables 5-8 show the results of similarities of fuzzy clusters shown in equation (10). Table 9 shows centers of fuzzy clusters. We capture objects which for the most part belong to a fixed cluster and calculate the centers of the objects for each fixed cluster. Figure 8 shows the visualization of this table. From this figure, we can see that cluster 1 shows subjects who mainly consider "taste" and "service", while those belonging to cluster 2 give higher scores for "taste", "cost", and "quantity". Cluster 3 shows subjects who give high scores for all variables especially for the "cost" and who do not visit the restaurant many times. Cluster 4 also has subjects who give high scores for all variables, but who have the opposite tendency to the cluster 3, in that they visit the restaurant many times. Figures 9 and 10 show degree of belongingness of subjects of male and female to fuzzy clusters, respectively. According to the values of degree, the subjects are sorted for each cluster. From these figures, there is a significant difference between male and female subjects for clusters 2 and 4. That is tendencies of these clusters are opposite to male and female subjects. From figure 8, the difference between clusters 2 and 4 is the value of frequency. That is female subjects do not visit restaurants while male students visit restaurants many times. Based on these features of clusters shown in figures 8-10, we can see the differences over the 4 industries through the obtained similarity of fuzzy clusters shown in tables 5-8.

Table 2 Attribute data

Subjects	Age	Sex	Belongs
1	24	1	1
⋮	⋮	⋮	⋮
46	21	2	1

[sex: 1:male, 2:female;] [belongs: 1:sports, 2:culture, 3:no belongs]

Table 3 Questionnaire data

Industries	Subjects	Taste	Cost	Frequency	Service	Quantity
1	1	3	4	2	4	2
	⋮	⋮	⋮	⋮	⋮	⋮
	46	4	3	3	5	3
2	1	3	2	1	3	1
	⋮	⋮	⋮	⋮	⋮	⋮
	46	4	5	4	4	4
3	1	3	3	2	3	2
	⋮	⋮	⋮	⋮	⋮	⋮
	46	3	3	1	3	4
4	1	3	3	0	4	2
	⋮	⋮	⋮	⋮	⋮	⋮
	46	5	5	2	4	2

Table 4 Result of fuzzy clustering

Industries	Subjects	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	1	41	29	26	4
	⋮	⋮	⋮	⋮	⋮
	46	9	10	23	58
2	1	35	43	13	10
	⋮	⋮	⋮	⋮	⋮
	46	1	4	61	34
3	1	65	31	3	1
	⋮	⋮	⋮	⋮	⋮
	46	8	81	9	2
4	1	80	15	4	1
	⋮	⋮	⋮	⋮	⋮
	46	18	17	53	12

Table 5 Similarity matrix of fuzzy clusters in industry 1

Fuzzy Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	14.3	11.6	9.4	5.6
Cluster 2	11.6	10.9	9.5	5.5
Cluster 3	9.4	9.5	9.6	5.1
Cluster 4	5.6	5.5	5.1	3.5

$(\times 10^5)$

From these tables, we can see that the self-similarity of cluster 1 is large for industries 1, 3, and 4 while the self- similarity of cluster 4 is large for industry 2. That is, for restaurants serving “hamburgers”, “noodles”, and “doughnuts”, subjects strongly consider taste and service, but do not frequently visit such restaurants. However, they consider almost all variables except cost for the restaurants serving “rice”, and they visit this type of restaurant many times. In industries 1, 3, and 4, clusters 1-3 are similar to each other when compared with cluster 4. However, in industry 2, there is an opposite situation, that is, clusters 2 and 4 are similar to each other. This means that similar subjects who give similar types of scores over the five variables belong to clusters 2 and 4 simultaneously for the industry 2. That is, subjects give scores which do not relate with frequency of visit for restaurants serving “rice”.

Table 6 Similarity matrix of fuzzy clusters in industry 2

Fuzzy Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	5.6	7.0	5.8	8.4
Cluster 2	7.0	9.2	7.7	10.7
Cluster 3	5.8	7.7	6.7	9.1
Cluster 4	8.4	10.7	9.1	13.1

$(\times 10^5)$

Table 7 Similarity matrix of fuzzy clusters in industry 3

Fuzzy Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	12.8	11.3	8.5	7.8
Cluster 2	11.3	10.5	7.9	7.4
Cluster 3	8.5	7.9	6.1	5.6
Cluster 4	7.8	7.4	5.6	5.5

$(\times 10^5)$

Table 8 Similarity matrix of fuzzy clusters in industry 4

Fuzzy Clusters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	10.4	10.1	7.8	9.2
Cluster 2	10.4	10.3	7.7	9.3
Cluster 3	7.8	7.7	6.3	7.3
Cluster 4	9.2	9.3	7.3	8.8

$(\times 10^5)$

Table 9 Centers of clusters

Fuzzy Clusters	Taste	Cost	Frequency	Service	Quantity
Cluster 1	3.54	2.69	1.33	3.40	1.83
Cluster 2	3.33	3.16	1.71	2.67	3.24
Cluster 3	3.57	4.51	2.57	3.62	3.73
Cluster 4	4.50	2.95	3.68	3.70	3.73

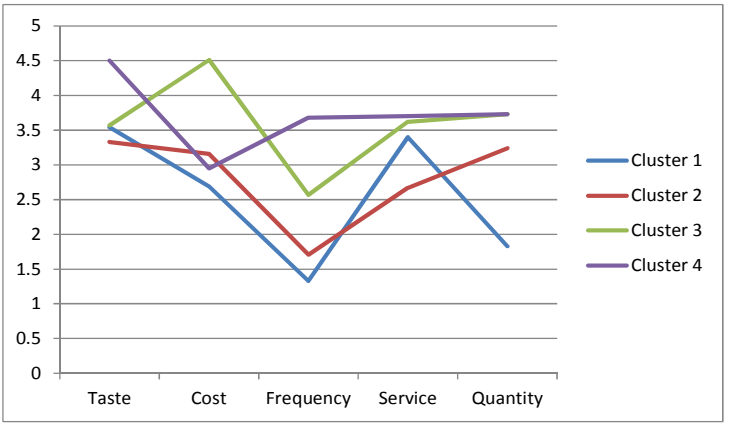


Fig. 8 Centers of clusters

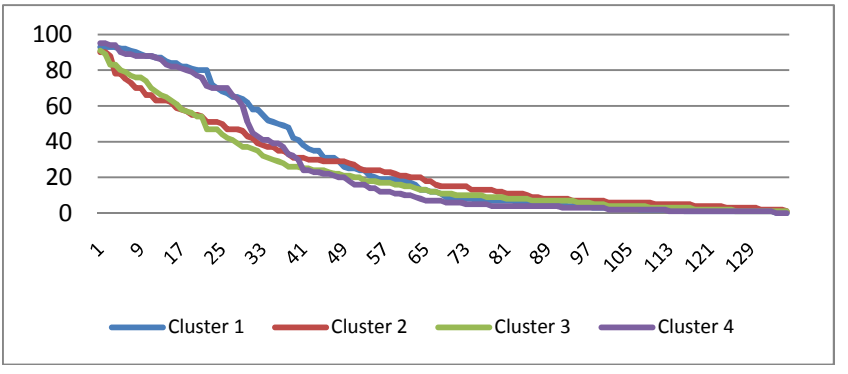


Fig. 9 Degree of belongingness of male subjects to clusters

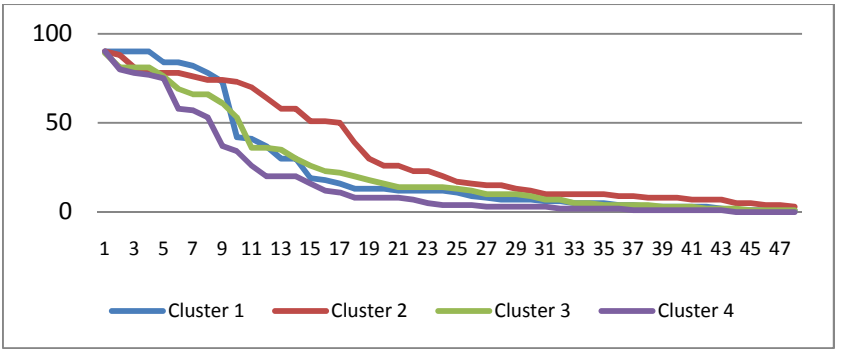


Fig. 10 Degree of belongingness of female subjects to clusters

6 Conclusions

This paper proposes a method to capture the similarity of fuzzy clusters which are a reduction of similarity of objects. The merit of this method is to obtain the similarity structure in reduced dimensional space spanned by the fuzzy clusters. Using the same fuzzy clusters, the changing situation of temporal data can be obtained under comparable results. Several numerical examples show the better performance of our proposed method.

References

1. Couper, M.P.: Designing Effective Web Surveys. Cambridge University Press, New York (2008)
2. Couper, M.P.: The Future of Modes of Data Collection. Public Opinion Quarterly, 75(5), 889–908 (2011)

3. Heiser, W.J., Groenen, P.J.F.: Cluster Differences Scaling with A Within-Clusters Loss Component and A Fuzzy Successive Approximation Strategy to Avoid Local Minima. Psychometrika 62(1), 63–83 (1997)
4. Kaufman, L., Rousseeuw, P.J.: Finding Groups in Data. Wiley (1990)
5. Manfreda, K.L., Bosnjak, M., Berzelak, J., Haas, I., Vehovar, V.: Web Sur-veys versus Other Survey Modes: A Meta-Analysis Comparing Response Rates. International Journal of Market Research 50(1), 79–104 (2008)
6. Millar, M.M., Dillman, D.A.: Improving Response to Web and Mixed-Mode Surveys. Public Opinion Quarterly 75(2), 249–269 (2011)
7. Sato, M., Sato, Y., Jain, L.C.: Fuzzy Clustering Models and Applications. Springer, Heidelberg (1997)
8. Sato-Ilic, M., Jain, L.C.: Innovations in Fuzzy Clustering. Springer, Heidelberg (2006)
9. Schönemannand, P.H., Wang, M.M.: An Individual Difference Model for the Multi-dimensional Analysis of Preference Data. Psychometrika 37(3), 275–309 (1978)
10. Valliant, R., Dever, J.A.: Estimating Propensity Adjustments for Volunteer Web Sur-veys. Sociological Methods & Research 40(1), 105–137 (2011)

A Study of Maslow's Hierarchy of Needs and Decision-Making

Hiroyuki Kishi, Koji Ito, Shin Sugiura, and Eizo Kinoshita

Abstract. This study attempts to clarify the psychological structures of people's values and their decision-making processes while validating the Analytic Hierarchy Process (AHP). Within this attempt, we apply Maslow's hierarchy of needs [1][2] to the framework of this structure, and we propose a view of decision-making that is composed of two systems. One system is value judgment pressure expressed by value judgment distance (ego-involvement scale), and the other system is decisions made from value selection pressure expressed by distance among basic needs. (basic needs [1][2]) This is a view of the basic structure of values and decision-making.

1 Introduction

This study comprises research related to landowners' structure of values and decision-making which has been positioned as progressive research that investigates succession awareness (consciousness) accepted by reconstruction of the house (residences) [3]. Focusing on the need for succession that was included in previous studies although it does not have a place in Maslow's hierarchy of five basic needs, this research has confirmed the existence of its existence using multivariate analyses such as AHP.

Moreover, in order to regard these studies as progressive, the clarification of the psychological structures by which people discover their values and make decisions has been made a primary topic.

Hiroyuki Kishi · Koji Ito · Shin Sugiura · Eizo Kinoshita
Meijo University, Nagoya, Japan
e-mail: h-design-lab@hi.enjoy.ne.jp, kohji@meijo-u.ac.jp,
shinsu@urban.meijo-u.ac.jp, kinoshit@urban.meijo-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 89–97.
springerlink.com

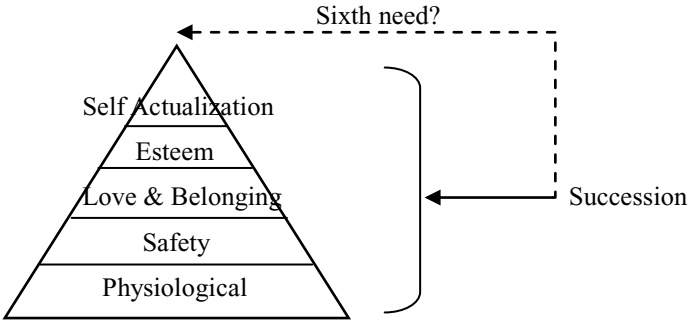


Fig. 1 Maslow's hierarchy of needs

2 Proof of Existence of the Need for Succession

We conducted a questionnaire survey to demonstrate the existence of the need for succession. By following the AHP and factor analysis data routes shown in Figure 2, we used a 2-way ANOVA to confirm a statistically significant difference in average scores of consciousness factors extracted from each of the value clusters (AHP outcome measure and alternatives) and factor analysis. [3]

The subject of AHP was values in house reconstruction, with two outcome measures: reconstruction that took the future usage by the next generation into consideration, and reconstruction that allows enjoyable living in the present. Both measures imply the intention for succession. Life activities during reconstruction based on Maslow's basic needs were used as alternative values. The survey was administered to 193 people, mainly consisting of average middle-aged and elderly individuals. With the exception of the eight people whose responses were invalid, we obtained 171 cases of data that can be used for AHP (paired comparisons with alternative: 57 items, C.I. < 0.15) and 161 cases of data for factor analysis.

3 Results of Analysis (Confirmation of Need for Succession) [3]

(1) We discovered the following value clusters from AHP:

a) Outcome measure scale: Values that consider the current generation (C'1), Values that consider the next generation (C'2), Values that consider both generations equally (C'3).
b) Alternative scale: Values emphasizing hierarchy (C1, C3), Values emphasizing equality hierarchy (C2).

(2) The shape of Maslow's hierarchy of needs were identified from the levels of importance of each alternative cluster in AHP (Figure 3).

(3) We obtained the following consciousness factors by factor analysis (principal factor analysis, promax rotation):

a) Residence lifestyle general evaluation scale: succession consciousness, general residence consciousness, reception of guests.
b) Current lifestyle evaluation scale: succession consciousness in stability orientation, self-esteem consciousness, family consciousness.
c) Residence reconstruction evaluation scale: Desires in reconstruction, succession orientation, self-esteem/self- actualization consciousness.

(4) We obtained the following results from the 2-way ANOVA:

a) C'1 and C'2 in AHP outcome measure scale clusters were statistically different at a 5% level in all current lifestyle consciousness factors (Figure 4 profile plot).
b) C'1 and C'2 in AHP outcome measure scale clusters were statistically different at a 5% level in all residence reconstruction consciousness factors (Figure 5 profile plot).
c) Although statistically significant differences were not seen from other clusters or consciousness factors, consciousness abstracted from average scale score trends and factor analysis expressed the possible existence of succession consciousness and succession orientation.

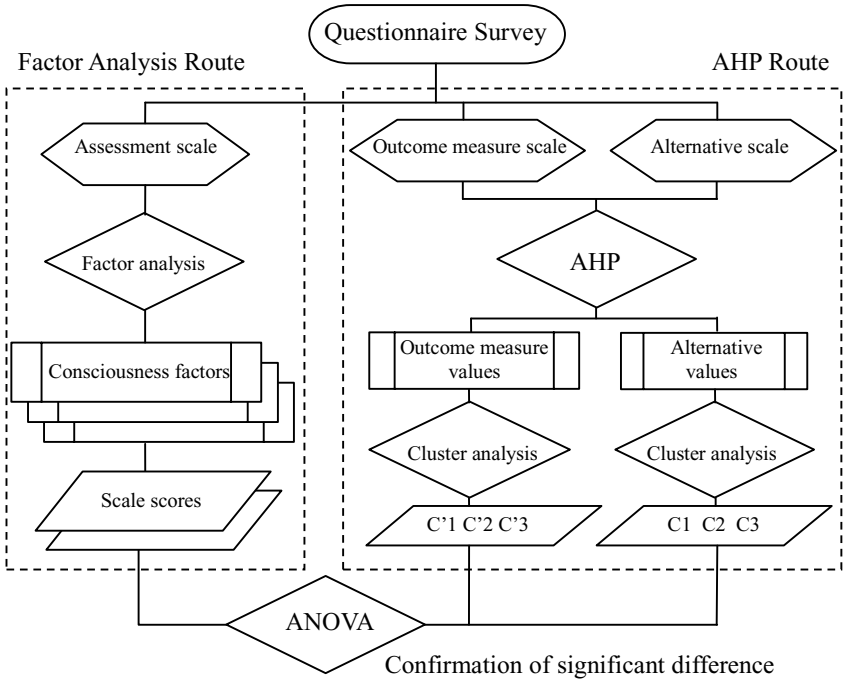


Fig. 2 Routes for discovering the need for succession

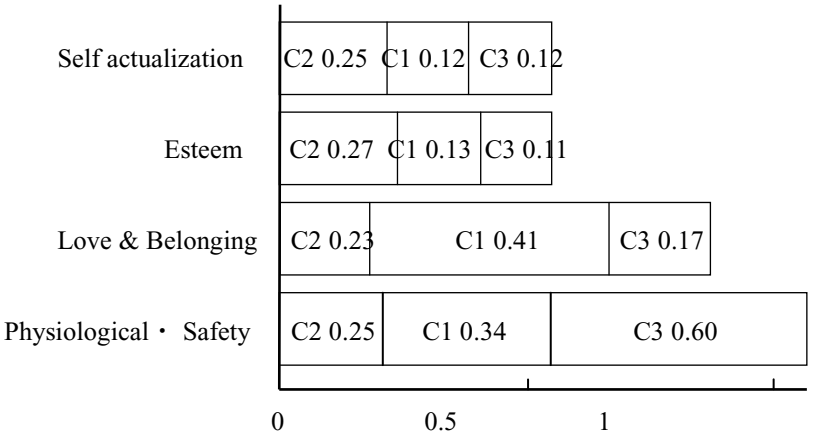


Fig. 3 Shape of Maslow's hierarchy as seen from levels of importance of alternative clusters

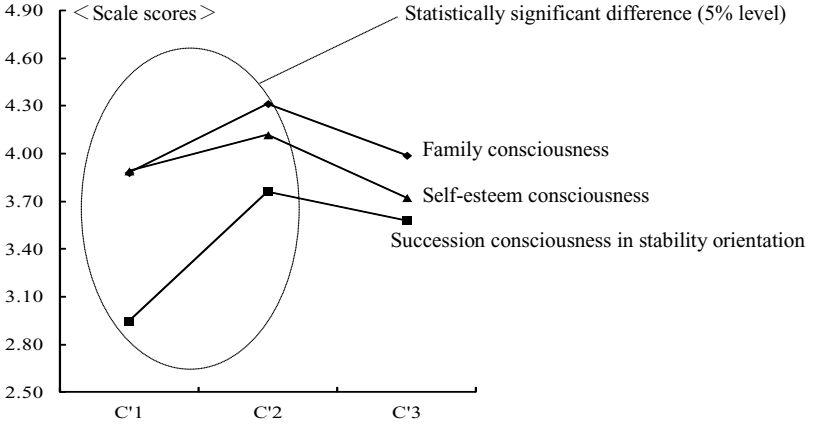


Fig. 4 Outcome measure scale clusters × Current lifestyle consciousness factors

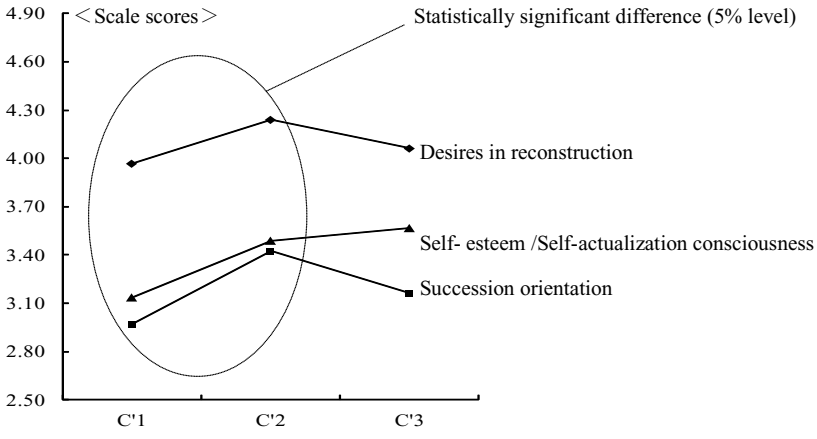


Fig. 5 Outcome measure clusters × Residence reconstruction consciousness factors

4 Basic Structure of Values and Decision-Making

From our results, we were able to theorize the existence of the need for succession in relationship to Maslow's hierarchy of basic needs by using AHP and factor analysis. Next, this study will investigate the basic structure of values and decision-making that can be considered as motivating factors that lead from need to action. Moreover, since Maslow's hierarchy of needs is essential when considering this basic structure, we have selected in advance value theories (key points) that are thought of as necessary. [1][2] Some of the key points include: (Point 1) We consider growth motivation to be the factor that leads to the upper levels in the hierarchy of needs. There are multiple determinants and varied levels of distance from basic needs and motivation for actions. (Point 2) Interests change when transitioning from old satisfaction to new satisfaction, which gives rise to change in a person's values. Good things that have already been obtained are taken for granted. (Point 3) There are low-level and high-level basic needs, and these needs exist as a hierarchy. (Point 4) People who have satisfied their high-level and low-level needs feel value in their high-level needs. Based on these value theories, this study proposes a basic structure of values and decision-making composed of two systems. These are,

(1) Decision-making processes determined by value judgment distance and value judgment pressure.

(2) Decision-making processes determined by distance among basic needs and value selection pressure.

What should be noted here is that although people value high-level needs, this does not imply that there is no value in low-level needs. As Maslow says, a change in people's values only causes them to take good things already obtained for granted. When both comparison targets are forcibly brought into one's

consciousness simultaneously, judging from the principles of the hierarchy, people discover importance in low-level values, that is, values relating to their own lives.

4.1 Psychological Distance

The value judgment distance and distance among basic needs discussed in this study represent psychological distance. Examples of this distance are shown in Table 1. The questions used to abstract succession factors were as follows: Q.a questions about you yourself, Q.b questions are about "If you were to…," and Q.c. questions about general questions.

These questions are structured to deepen one's involvement in oneself (henceforth considered psychologically as ego-involvement), and we can compare the tendency for the average scale scores to occur. We render this as a value judgment score that represents a concept of distance.

(The factor analysis scale score is the average scale score of abstracted consciousness factors, determined by respondents' ratings on the questionnaire as follows: "Disagree" = 1 point, "Somewhat disagree" = 2 points, "Neither agree nor disagree" = 3 points, "Somewhat agree" = 4 points, and "Agree" = 5 points.) On the other hand, distance among basic needs is represented as a concept of distance between each successive need on Maslow's hierarchy; the more the need relates to the basic needs of life, the shorter the distance.

Table 1 Factor Analysis Scale Scores for Identifying Need for Succession (from [3])

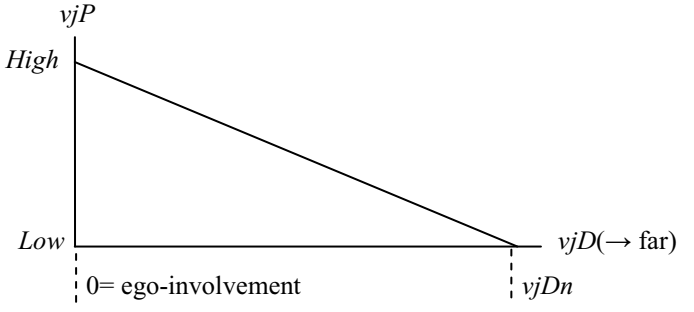
Question		Q.a	Q.b	Q.c
Consciousness factors abstracted from factor analysis		Succession consciousness in stability orientation	Succession orientation	Succession consciousness
AHP Outcome measure clusters	C'1	2.94	2.97	2.85
	C'2	3.76	3.43	3.12
	C'3	3.58	3.16	3.01
AHP Alternative clusters	C1	3.37	3.30	2.89
	C2	3.25	3.11	2.86
	C3	3.67	3.00	2.87

Q.a: Questions about your current lifestyle
Q.b: Questions about how you would reconstruct your current residence
Q.c.: Questions about general lifestyle and residence

4.2 Value Judgment Distance and Value Judgment Pressure

Value judgment pressure does not represent the value itself, but rather is a conceptual scale determined based on value judgment distance. The more ego-involvement there is, the greater the pressure becomes, and this is expressed

highly as a value held by the subject. This is similar to the importance calculated from AHP and the factor analysis scale score. If the value judgment distance for each comparison target is the same, a selection of "both are equally important" can be expected on the AHP paired comparison scale, and a selection of "neither agree nor disagree" (3 points on a 5-point scale) can be expected on the factor analysis scale. This leads to a negative correlation between value judgment distance and value judgment pressure as shown in Figure 6.

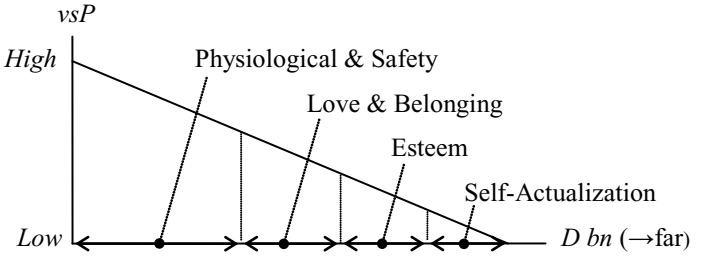


vjP = 1/vjD (*vjP* = value judgment pressure, *vjD* = value judgment distance)

Fig. 6 Relationship between Value Judgment Distance and Value Judgment Pressure

4.3 Distance among Basic Needs and Value Selection Pressure

Value selection pressure is similar in nature to value judgment pressure. Value selection pressure also does not represent the value itself, but rather is a scale of a concept determined from the distance among basic needs. The lower the need in Maslow's hierarchy (physiological and safety needs), the greater the pressure becomes, and this is expressed as a high value held by the subject. (highly necessary for living) Distance among basic needs and value selection pressure form a negative correlation as shown in Figure 7.



vsP = 1/ D bn (*vsP* = value selection pressure, *Dbn* = distance among basic needs)

Fig. 7 Relationship between Distance among Basic Needs and Value Selection Pressure

5 Conclusion

This study has proposed a basic structure in which a person's discovery of values and their decision-making actions are formed by value judgment distance and value judgment pressure and by distance among basic needs and value selection pressure. In the arena of decision-making models, Kinoshita [4] [5] has proposed a "decided model, establishing model, and deciding model," as shown in Figure 8, and he has defined mathematical properties for AHP. In the descriptive model of decision making, Kinoshita used the geometric mean method to minimize the difference between measured value and true value and discussed the necessity of continuity in consensus decision-making. In the normative model, he used the eigenvector method to solve the equilibrium point of the normative opinion (action) and discussed the necessity of independent value standards. Moreover, Kinoshita considers the AHP-establishing model as a concept that combines independence and continuity. If we adapt these concepts to the basic structure put forth in this study, decision-making expressed by value judgment distance and value judgment pressure corresponds to Kinoshita's decided model, and decision-making expressed by distance among basic needs and value selection pressure can be regarded as analogous to his deciding model.

Lastly, the issue of independence and dependency in AHP paired comparison is thought to be related to the distance scale and pressure difference scale in this study's proposed basic structure, and depending on the development of psychological standards, there is a possibility that a solution will be conceived.

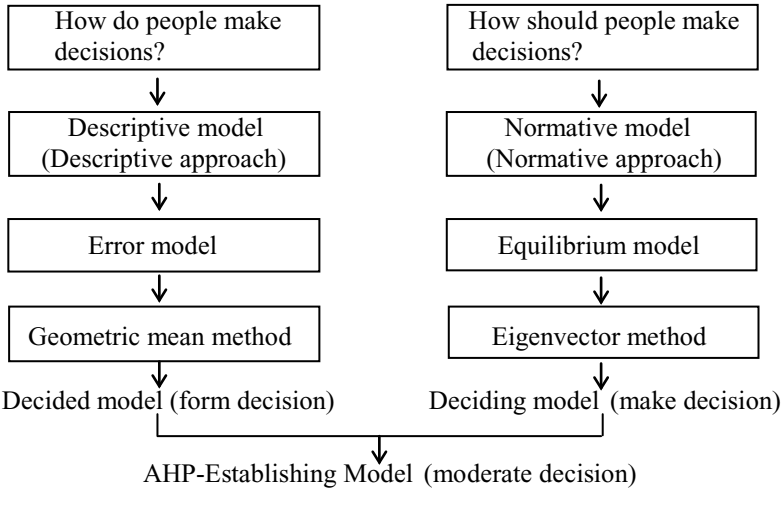


Fig. 8 Kinoshita's model (Decided model, Establishing model, Deciding model) (Quote from [4][5])

References

1. Maslow, A.H.: Toward a psychology of being, 2nd edn. Van Nostrand Reinhold Company Inc., New York (1968); Ueda, Y.(trans.): 2nd ed. Seishin Shobo (1998)
2. Maslow, A.H.: Motivation and personality, 2nd edn. Harper & Row Publishers, Inc. (1970); Oguchi, T.(trans.): Sangyo Nohritu Daigaku Publishing Division (1987)
3. Kishi, H.: A study of succession awareness accepted by reconstruction of the house. Meijo University Master's Thesis (2010)
4. Kinoshita, E.: Strategic model of The Art of War by Sun Tzu. Omsha (2006)
5. Kinoshita, E.: Service science. Service value measurement method -Cases of 10. Kindai Kagakusha (2009)

Adjustment Mechanism of Concurrent Convergence Method Models

Shin Sugiura and Eizo Kinoshita

Abstract. The paper describes how the adjustment mechanism of Concurrent Convergence Method (CCM) models work. In doing so, the paper examines the following three models: Weight Concurrent Convergence Method (WCCM), a model for adjusting the instability of weight of criteria; Evaluation Concurrent Convergence Method (ECCM), a model for adjusting the instability of evaluation values; and Total Evaluation Concurrent Convergence method (TCCM), a model for resolving the instability of evaluation values, caused by the instability of weight of criteria, through convergence calculation of total evaluation value. The authors believe that it may be an unprecedented attempt to analyze the three above-mentioned models, or WCCM, ECCM and TCCM, all at the same time, and that their comprehensive analysis of CCM models could contribute significantly to this area of study.

1 Introduction

Conventional Analytical Hierarchy Process (AHP) lacks a method to adjust differences among weight of criteria, as well as the instability of evaluation values. CCM models, however, are capable of coping with those problems. Among those models, Weight Concurrent Convergence Method (WCCM), proposed by Kinoshita and Nakanishi, for instance, is a method for adjusting differences among weight of criteria, which vary depending on respective dominant alternatives. Evaluation Concurrent Convergence Method (ECCM) and Total Evaluation Concurrent Convergence method (TCCM), both proposed by Sugiura and Kinoshita, are also those with a capacity for adjusting the instability of evaluation values. The ECCM is a method for adjusting differences among several evaluation values of an alternative. The TCCM is a method to acquire total evaluation value based on Weight Concurrent Convergence Method (WCCM), by applying the calculation method of the ECCM.

Shin Sugiura · Eizo Kinoshita
Meijo University, Kani, Japan
e-mail: shinsu@urban.meijo-u.ac.jp, kinoshit@urban.meijo-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 99–107.
springerlink.com

With the aim of achieving a comprehensive understanding of the adjustment mechanism of CCM models, the authors analyze and describe each of those methods in this paper, in the following order: The WCCM is explained in Chapter 2, which is followed by the analysis of the ECCM and the TCCM in Chapter 3, and finally in Chapter 4, a conclusion.

2 Dominant AHP and Weight Concurrent Convergence Method

2.1 Dominant AHP

In this section, the paper explains Dominant AHP[1] and Weight Concurrent Convergence Method (WCCM)[2], both proposed by Kinoshita and Nakanishi.

Dominant AHP is a method which focuses on a particular alternative, and conducts evaluation based on this alternative. In this method, such an alternative which is selected as a benchmark, is called "a dominant alternative," while the other alternatives are called "dependent alternatives." Here, the authors consider a case where three alternatives, X, W, and Z are selected as "dominant alternatives," and their respective weight of criteria are denoted by W_X, W_Y, W_Z. Then a matrix comprising those elements is expressed as below.

$$W=\left(W_X,W_Y,W_Z\right) \tag{1}$$

Evaluation value matrix M of alternatives under criteria is expressed by Formula (2).

$$M=\begin{bmatrix} a_{X\,\mathrm{I}} & a_{X\,\mathrm{II}} \\ a_{Y\,\mathrm{I}} & a_{Y\,\mathrm{II}} \\ a_{Z\,\mathrm{I}} & a_{Z\,\mathrm{II}} \end{bmatrix} \tag{2}$$

Matrix M_i is defined as is shown in Formula (3).

$$M_i=M\begin{bmatrix} 1/a_{i\mathrm{I}} & 0 \\ 0 & 1/a_{i\,\mathrm{II}} \end{bmatrix}=MA_i^{-1} \tag{3}$$

It should be noted that $A_i=\begin{bmatrix} a_{i\,\mathrm{I}} & 0 \\ 0 & a_{i\,\mathrm{II}} \end{bmatrix}$ (i=X,Y,Z).

Dominant AHP is a method where a particular alternative is selected and made a criterion to evaluate other alternatives.

When Alternative X is selected as a dominant alternative, weight of criteria in terms of Alternative X is expressed by Formula (4).

$$A_X A_X{}^{-1} W_X = W_X \tag{4}$$

Evaluation values under criteria is expressed by Formula (5), in which the evaluation value of Alternative X is normalized to 1.

$$M_X = MA_X{}^{-1} \tag{5}$$

Total evaluation value is expressed by Formula (6).

$$M_X \left(A_X A_X{}^{-1} W_X \right) = MA_X{}^{-1} W_X \tag{6}$$

Estimate values of weight of criteria in terms of Alternative Y can be expressed by $A_Y A_X^{-1} W_X$, bringing total evaluation value to be expressed by Formula (7).

$$M_Y \left(A_Y A_X{}^{-1} W_X \right) = MA_X{}^{-1} W_X \tag{7}$$

By the same token, estimate values of weight of criteria in terms of Alternative Z can be expressed by $A_Z A_X^{-1} W_X$, bringing total evaluation value to be expressed by Formula (8).

$$M_Z \left(A_Z A_X{}^{-1} W_X \right) = MA_X{}^{-1} W_X \tag{8}$$

Thus, Formula (6), (7) and (8) prove that Dominant AHP owns different total evaluation value depending on each alternative, and that the total evaluation value of Alternative X, and that of Alternative Y as well as of Alternative Z are all identical to one another.

2.2 Weight Concurrent Convergence Method

The authors now describe Weight Concurrent Convergence Method (WCCM) [2][3]. One of the main characteristics of the WCCM is that there are several dominant alternatives in Dominant AHP, and that a criterion varies depending on respective dominant alternatives.

First, the authors examine a case where not only Alternative X, but also Alternative Y, are selected as dominant alternatives. Total evaluation value is expressed by Formula (9), just like in Formula (6).

$$M_Y \left(A_Y A_Y{}^{-1} W_Y \right) = MA_Y{}^{-1} W_Y \tag{9}$$

Estimate values of weight of criteria in terms of Alternative X can be expressed by $A_X A_Y^{-1} W_Y$. When $A_X A_Y^{-1} W_Y$ equals with Weight W_X , which is obtained in terms of a case where Alternative X is selected as a dominant alternative, total evaluation value is expressed by Formula (10).

$$M_Y\left(A_Y A_Y^{-1} W_Y\right) = M_Y\left(A_Y A_X^{-1} W_X\right) = M A_X^{-1} W_X \tag{10}$$

It demonstrates that the outcome is identical with the one obtained when Alternative X is Dominant Alternative X.

In reality, however, W_X and estimate value $A_X A_Y^{-1} W_Y$ rarely conform to each other, and it often results in developing differences. Thus, Kinoshita and Nakanishi developed the WCCM to adjust such "differences" of criteria.

In the WCCM, an adjusted value of weight is obtained by calculating the mean of adjusted values in terms of several dominant alternatives, which are given at the outset. When e represents a vector where each of its all the elements is 1, while T represents disposition, and when all the alternatives are dominant alternatives, adjusted value $\tilde{W}_X$ of weight of criteria in terms of Dominant Alternative X is expressed by Formula (11).

$$\tilde{W}_X = \frac{1}{3}\left(\frac{A_X A_X^{-1} W_X}{e^T A_X A_X^{-1} W_X} + \frac{A_X A_Y^{-1} W_Y}{e^T A_X A_Y^{-1} W_Y} + \frac{A_X A_Z^{-1} W_Z}{e^T A_X A_Z^{-1} W_Z} \right) \tag{11}$$

By the same token, adjusted values $\tilde{W}_Y$ and $\tilde{W}_Z$, of weight of criteria in terms of Dominant Alternative Y and Z, respectively, are shown as below.

$$\tilde{W}_Y = \frac{1}{3}\left(\frac{A_Y A_X^{-1} W_X}{e^T A_Y A_X^{-1} W_X} + \frac{A_Y A_Y^{-1} W_Y}{e^T A_Y A_Y^{-1} W_Y} + \frac{A_Y A_Z^{-1} W_Z}{e^T A_Y A_Z^{-1} W_Z} \right) \tag{12}$$

$$\tilde{W}_Z = \frac{1}{3}\left(\frac{A_Z A_X^{-1} W_X}{e^T A_Z A_X^{-1} W_X} + \frac{A_Z A_Y^{-1} W_Y}{e^T A_Z A_Y^{-1} W_Y} + \frac{A_Z A_Z^{-1} W_Z}{e^T A_Z A_Z^{-1} W_Z} \right) \tag{13}$$

When replacing $\tilde{W}_X$, $\tilde{W}_Y$ and $\tilde{W}_Z$ with W_X, W_Y, and W_Z, respectively, and repeating the same calculation stated above, the result converges and is expressed by W_i^*(i=X, Y, Z), which denotes weight of criteria. The convergent property of the above-mentioned result, derived from weight of criteria, is substantiated by Reference [3]. Then product $M_i W_i^*$, obtained by multiplying Matrix M_i by converged weight W_i^*, leads to evaluation value E_i of respective alternatives, as is shown in Formula (14).

$$E_i = M_i W_i^* \tag{14}$$

When normalizing Formula (14) so that the total shall be 1, it results in bringing out the same vector in every case, which represents total evaluation value.

Concerning the WCCM, it should also be noted that in geometric mean CCM, proposed by Ohya and Kinoshita [6], geometric mean is utilized for adjusting weight, based on an error model. This method, by using geometric mean, takes out

the need to repeat convergence calculation of Formulas (11) through (13), making a one-shot adjustment possible.

3 Evaluation and Total Evaluation Concurrent Convergence Method

3.1 Evaluation Concurrent Convergence Method

Evaluation Concurrent Convergence Method (ECCM) is a model proposed by Sugiura and Kinoshita [4]. Conventionally in AHP, including Dominant AHP and the WCCM, differences can emerge only in the area of weight of criteria, and no discussion has been made concerning the evaluation value itself. In reality, however, differences can develop among evaluation values, before weight differences come into being. Thus when there are several evaluation values, caused by differences among evaluation values, due to the existence of several selectors and ambiguous factors, it is necessary to have those values to be converged into a single value, because Dominant AHP and/or the WCCM is not applicable to this case.

The authors also note that when AHP is applied in the process of building a consensus in a group, there could often be several evaluation values. The ECCM, developed to deal with such a case, is a method to make adjustment have diversified values to be converged into a single value.

The ECCM has been proved effective in converging diversified evaluation values in the following circumstances: At a time when several decision makers engage in decision making and consensus building; and when there are several evaluation values, prompted by additional information which obliges people to review the adequacy of a decision they have made earlier.

Here, the authors explain this process by using n, which represents the number of evaluators, and m which denotes the number of criteria, as well as l which represents the number of alternatives [5]. Evaluation values by Evaluator i would be expressed by Formula (15).

$$M^i = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{lm} \end{bmatrix} \tag{15}$$

Next, the authors demonstrates Formula (16). In this formula, the product, obtained by multiplying the evaluation values by Evaluator i by a matrix in which reciprocals of the evaluation values of Alternative j are placed as diagonal elements, denotes the evaluation values in terms of Alternative j concerning all the criteria of Evaluator i. Just like in Formula (3) of the WCCM, in this formula, a particular alternative is selected and compared with other alternatives, by setting the evaluation value of the selected alternative at 1.

$$M_j^i = M^i \cdot \begin{bmatrix} 1/a_{j1} & 0 & \cdots & 0 \\ 0 & 1/a_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/a_{jm} \end{bmatrix} \tag{16}$$

This means that the evaluation value of Alternative *j* is set at 1 in terms of all the criteria. The authors then calculate the average of M_j and make it an initial value, with the number of evaluators being denoted by *n*.

$$M_j^{(1)} = \frac{1}{n}\sum_{i=1}^{n} M_j^i = \begin{bmatrix} a_{11}^j & \cdots & a_{1m}^j \\ \vdots & & \vdots \\ a_{l1}^j & \cdots & a_{lm}^j \end{bmatrix} \tag{17}$$

Concerning the transformation of Formula (17), it is possible to cause differences among weight of respective evaluators and decision makers, by applying weighted average method to the evaluation values of Evaluator *i*.

Formula (17) denotes the evaluation value obtained in terms of Alternative j as a criterion. In this formula, the evaluation value of Alternative j always becomes 1. That is to say, the formula expresses the relative evaluation ratio between Alternative j and the other alternatives. Thus, when multiplying the evaluation value in terms of Alternative *k* by a matrix, in which reciprocals of the evaluation values of Alternative *j* are placed as diagonal elements, it is able to newly derive the evaluation value in terms of Alternative *j* from the evaluation value in terms of Alternative *k*.

$$M_{kj} = M_k^{(1)} \cdot \begin{bmatrix} 1/a_{k1}^j & 0 & \cdots & 0 \\ 0 & 1/a_{k2}^j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/a_{km}^j \end{bmatrix} \tag{18}$$

In Formula (19), total number of alternatives is denoted by *l*, and second-generation values can be obtained by calculating the mean. In other words, the authors compare the evaluation value of a certain alternative with that of another alternative in the form of ratio, then calculate the mean of the evaluation values of the selected alternative as well as the other alternatives, and set it as the evaluation value of a newly selected alternative.

$$M_j^{(2)} = \frac{1}{l}\sum_{k=1}^{l} M_{kj} \tag{19}$$

By substituting the value derived from Formula (19) into Formula (18), and by repeating the calculation, it converges with $M_j^{(t)}=M_j^{(t-1)}$. In the ECCM, the geometric mean can be utilized to achieve the convergence.

Because the converged evaluation value signifies the evaluation value in terms of Alternative *j*, a single total evaluation value can be obtained by normalizing $M_j^{(t)}$ so that the total shall be 1.

The difference between the WCCM and the ECCM lies in the fact that the former is a method for adjusting differences among unstable criteria, while the latter is a method for adjusting differences among unstable evaluation values.

3.2 Total Evaluation Concurrent Convergence Method

Total Evaluation Concurrent Convergence Method (TCCM) is a method to acquire total evaluation value based on the WCCM, by applying the calculation method of the ECCM. Next, the paper explains the TCCM, proposed by Sugiura and Kinoshita[5].

In AHP, it aims at obtaining total evaluation value by multiplying the evaluation value of respective alternatives by a vector which represents weight of respective criteria. That is to say, when *M* represents evaluation values, and *W* represents weight of criteria, total evaluation value *E* is expressed by Formula (20).

$$E \equiv M \cdot W \tag{20}$$

Suppose that there is an issue which requires decision-making. When *n* represents criteria, and *m* represents the number of alternatives, the evaluation value of an alternative under respective criteria is expressed as below.

$$M = \begin{bmatrix} U_{11} & \cdots & U_{1j} & \cdots & U_{1n} \\ \vdots & & \vdots & & \vdots \\ U_{i1} & \cdots & U_{ij} & \cdots & U_{in} \\ \vdots & & \vdots & & \vdots \\ U_{m1} & \cdots & U_{mj} & \cdots & U_{mn} \end{bmatrix} \tag{21}$$

Then, a weight vector of criteria for respective alternatives is expressed by Formula (22), just like in the WCCM.

$$W = \left(W_1, \cdots, W_j, \cdots, W_m\right) \tag{22}$$

In Formula (22), W_j signifies weight of criteria in terms of Alternative *j*. And because W_j, or weight of criteria in terms of Alternative *j*, includes the weights equivalent of *n* (*n* represents the number of criteria), it can be expressed as below.

$$W_j = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \tag{23}$$

However, because overall weight of criteria becomes 1, it turns out that $\Sigma w_k = 1$.

$$M^j = M \cdot \begin{bmatrix} 1/U_{j1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/U_{jn} \end{bmatrix} \tag{24}$$

Next, in Formula (24), the evaluation value in terms of Alternative *j* as a new criterion, is derived from the evaluation value of an alternative given at the outset. In other words, the calculation is conducted in a manner as to bring the evaluation value of Alternative *j* to 1.

Thus, an initial value of Formula (25) is obtained.

$$E_1^j = M^j \cdot W_j = M \cdot \begin{bmatrix} 1/U_{j1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/U_{jn} \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} u_1^j \\ \vdots \\ u_m^j \end{bmatrix} \tag{25}$$

When setting the outcome of Formula (25) as an initial value, its second generation turns out to be a column vector, which conforms to the principles of the ECCM, and Formula (26) is obtained as below.

$$E_2^j = \frac{1}{m} \sum_{k=1}^{m} \frac{E_1^j}{u_k^j} \tag{26}$$

By the same token, Formula (27) is obtained.

$$E_{n+1}^j = \frac{1}{m} \sum_{k=1}^{m} \frac{E_n^j}{u_k^j} \tag{27}$$

By repeating this manner, the calculation converges with $E_{n+1}^j = E_n^j$. Then, Formula (28) is derived from E_{n+1}^j.

$$E_{n+1}^j = \begin{bmatrix} u_1^j \\ \vdots \\ u_m^j \end{bmatrix} \tag{28}$$

Finally, total evaluation value denoted by E is obtained by normalizing Formula (28).

$$E = \frac{u_k^j}{\sum_{k=}^{m} u_k^j} \tag{29}$$

In the TCCM, the number of evaluation values of respective alternatives, with each evaluation value being set at 1, corresponds to m, which represents the total number of alternatives. It is possible to obtain a single total evaluation value through a normalization process. One of the main characteristics of the TCCM is that it focuses on the determination of evaluation itself, unlike the WCCM which focuses on weight of criteria.

4 Conclusion

The paper demonstrates a comprehensive analysis of the adjustment mechanism of CCM models. While the WCCM is a model for adjusting the instability of weight of criteria, the ECCM is a model for adjusting the instability of evaluation values. Finally, the TCCM is a model for resolving the instability of evaluation values caused by the instability of weight of criteria through convergence calculation of total evaluation value.

The authors believe that it is a vital future task to apply diversified CCM models to various issues, and to prove the effectiveness of the CCM as a mathematical decision-making model.

References

1. Kinoshita, E., Nakanishi, M.: Proposal of New AHP Model in light of Dominant Relationship among Alternatives. Journal of Operations Research Society of Japan 42(2), 180–197 (1999)
2. Kinoshita, E., Nakanishi, M.: A Proposal of CCM as a Processing Technique for Additional Data in the Dominant AHP. Journal of Infrastructure Planning and Management IV-42(611), 13–19 (1999) (in Japanese)
3. Kinoshita, E., Sekitani, K., Shi, J.: Mathematical Properties of Dominant AHP and Concurrent Convergence Method. Journal of Operations Research Society of Japan 45(2), 198–213 (2002)
4. Sugiura, S., Kinoshita, E.: Proposal of Evaluation Concurrent Con-vergence Method. Journal of Infrastructure Planning 21, 33–44 (2004) (in Japanese)
5. Sugiura, S., Kinoshita, E.: Proposal of Total Evaluation Concurrent Convergence Method. Journal of Infrastructure Planning 22, 39–46 (2005) (in Japanese)
6. Ohya, T., Kinoshita, E.: The Geometric Mean Concurrent Convergence Method. In: Proceeding of the International Symposium on the Analytic Hierarchy Process 2009 (2009), http://www.isahp.org/2009Proceedings/index.htm

An Alternative View of Uncertainty in Bilateral Bargaining Models with Incomplete Information

Javier González and Vitaliy V. Kalashnikov

Abstract. This paper seeks to analyze the implications of relaxing the assumption of rationality by using incomplete information. Specifically, we want to analyze the implications on the uniqueness and existence of equilibrium when individuals form expectations with different distributions of all possible states. This is done by postulating an alternative view of uncertainty in decision making in economic interactions. Thus an alternative model of bilateral bargaining with incomplete information is formulated.

1 Introduction

Over the last 60 years game theory has become an important tool in economic modeling of strategic decision making in interactive situations. Particularly, game theoretic models have provided a solid framework for the study of strategic interaction previously lacking in the field of economics. The introduction of a higher level of mathematical precision and sophistication has come, however, at the expense of accuracy of its predictions and consistency with observed behavior. That is, fundamental assumptions have become more restrictive in the development of oversimplified models trying to predict behavior. One such fundamental assumption in game theory is rational behavior by the decision-maker, or *rationality* [8]. Within the context of game theory rationality refers to the ability of an individual to consistently make decisions that maximize that individual's expected payoff given varying levels of knowledge. In the absence of *'uncertainty'* rationality implies an individual's full knowledge (i.e., complete information) of the structure of the game, cognitive ability to weigh different choices and deliberately make optimal decisions. In the presence of uncertainty about the game structure (i.e., incomplete information) rationality

Javier González · Vitaliy V. Kalashnikov
Graduate School of Economics, UANL, Av. Lázaro Cárdenas S/N,
Monterrey, N.L., Mexico 64890
e-mail: javier.gonzalez.g@gmail.com, kalashnikov_de@yahoo.de

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 109–118.
springerlink.com

implies an individual's ability to choose actions that maximize expected payoffs given full knowledge of all possible games and their distribution. It is the assumption of rational behavior when players are uncertain about the probability distribution their opponents assign to the different 'states' of the game that we seek to relax. Then implications of relaxing said assumption on the existence and uniqueness of a Nash equilibrium are analyzed. Our goal is to introduce this alternative view of uncertainty in decision making in economic interactions with the purpose of formulating an alternative model of decision making in bilateral bargaining games with incomplete information.

Under the assumption of (super) rationality in the absence of uncertainty, in 1950 John Nash [6] developed an axiomatic framework to form an equilibrium solution to the bilateral bargaining problem. While valid under specific complete information conditions, Nash's bargaing model appears to be inconsistent with observed behavior when players are faced with incomplete information. Looking to overcome Nash's model inconsistencies, Jonh Harsanyi [2] formulated a model that accounts for players with less-than-complete information about the structure of the game. In the presence of (unmeasurable) uncertainty of the structure of bilateral bargaining game, Harsanyi assumes players will form a commonly known set of possible game structures and a corresponding probability distribution (i.e., imperfect information). That is, players will transform a game of unknown possible actions and outcomes into a game where the distribution of all possible actions and outcomes is commonly known. Thus, players can form an expectation of a set of possible actions that satisfies the conditions for a Nash Equilibrium.

While Harsanyi's model holds under common knowledge of different sets of actions and outcomes with a common probability distribution, it appears to be inconsistent with observed behavior once players' are endowed with different probability distributions. The objective of the present study is to analyze the implications of introducing unmeasurable uncertainty [4] (Knightian Uncertainty) into the existing framework of bilateral bargaining models. Specifically, we model bilateral bargaining where players assign different probabilities over all possible actions and outcomes leading to a disagreement that is inconsistent with theoretical predictions. Thus providing the foundation for a model of bilateral bargaining with incomplete information under unmeasurable uncertainty.

We begin our task in section 2 by describing the general bilateral bargaining model with incomplete information, and assuming rational behavior. We assume each player has full knowledge of all the possible states of the game, as well as the probability of each state. Thus we rewrite the game of incomplete information as a game of *imperfect* information. Next in section 3 we relax the assumption of rationality, introducing uncertainty of the probability a player assigns over all the possible states of the game. Specifically, we allow each player to have a different (subjective) probability distribution over all possible states. The implications of this are examined conceptually, then illustrated by an example. Finally, conclusions are provided in section 4.

2 Measurable Uncertainty and Mutual Consistency of Beliefs in Bilateral Bargaining

Under Nash's axiomatic model [6] with complete information for a bilateral bargaining game with a solution of the form

$$f^N(S,d) = \underset{(d_1,d_2)\leq(x_1,x_2)\in S}{\operatorname{argmax}} (x_1-d_1)(x_2-d_2)$$

subject to a bargaining set

$$S(d) = \left\{(x_1,x_2)\in S \,|\, x_1+x_2\leq c, x_i\geq d_i, S\subset R^2\right\}$$

where x_i is the payoff demand for player i within the agreement convex space S. This set is bounded below by a disagreement payoff d_i and it is assumed that both players have complete knowledge of each other's possible actions. To take account of incomplete information in games, Harsanyi [2] proposed a model that transforms a game of incomplete information into a game of imperfect information. Under the framework of incomplete information with measurable uncertainty proposed by Harsanyi, the information of the different sets of actions and payoffs are associated with a distribution of 'player types' with given probabilities. Thus, in the context of a bilateral bargaining game with a normal form for each player $i=1,2$ with an arbitrary number of types k, a vector of types for player i is represented by $c_i = \left\{c_i^1, c_i^2, \ldots, c_i^k\right\}$ which is used to normalize strategies under complete information for every possible k type such that $s_i \mapsto s_i^* = s_i^*(c_i) = \left\{s_i^*\left(c_i^1\right), s_i^*\left(c_i^2\right), \ldots, s_i^*\left(c_i^k\right)\right\}$, and each player i has a vector of subjective probabilities (probability distribution) of all the possible types for his opponent given his own, such that $\bar{P}_i(c_j) = P_i(c_j|\,c_i)$ for $j=1,2$, $j\neq i$. Let S_i be the set of all possible strategies for all k types for player $i=1,2$, $C_i=\{c_i\}$ the set of all vector types. Let the game of incomplete information be defined as $G=\{S_1,S_2,C_1,C_2,V_1,V_2,P_1,P_2\}$ where $V_i:[S_1\times S_2\times C_1\times C_2]\to R$ represents player i's payoff function.

Under the central assumption that the vectors of subjective probabilities of both players are equal, i.e., $\bar{P}_1(c_2)=\bar{P}_2(c_1): P_i(c_j|\,c_i)=P^*(c_j|\,c_i)$ for $i,j\in\{1,2\}$, $i\neq j$, Harsanyi's model claims that it is possible to derive a commonly known objective probability distribution such that

$$P^*(c_j|\,c_i) = \frac{P^*(c_i,c_j)}{\int_{C_j} d_{(c_j)} P^*(c_i,c_j)} \Rightarrow P^*(c_i,c_j) = P^*(c) = P_i(c_i,c_j)\int_{C_j} d_{(c_j)} P^*(c_i,c_j) \tag{1}$$

Using player i's normalized strategies, $s_i^*(c_i)$, normalized payoffs are derived, $x_i = V_i\left(s_1^*(c_1), s_2^*(c_2), c_1, c_2\right)$, which in turn provides a weighted sum of the normalized strategies that are used to represent player i's expected payoffs in the normal form of the game:

$$Ex_i = W_i(s_1^*, s_2^*) = \int_C V_i(s_1^*, s_2^*)\, d_{(c)} P^*(c)$$

which can be modified to account for player i's expected payoffs given his own type

$$E(x_i|c_i) = Z_i(s_1^*, s_2^*|c_i) = \int_{C_j} V_i(s_1^*, s_2^*, c_1, c_2)\, d_{(c_j)} P^*(c_j|c_i)$$

Assuming that for player i's normalized strategy, s_i^*, a specific vector of types, c_i^0, there exists a corresponding optimal ordinary strategy $s_i = s_i^*(c_i)$ that maximizes the expected payoff, $E\left(x_i|c_i^0\right) = Z_i\left(s_1^*, s_2^*|c_i^0\right)$ such that s_i^* is the best response for c_i^0 given his opponent's strategy s_j^*, then for s_i^* being a uniformly best response for c_i except for a small subset of types c_i^* – with probability nearly zero – $\left\{s_i^* \cup s_j^*\right\}$, $i, j \in \{1,2\}, i \neq j$ gives us an equilibrium point. This implies that in order to have an equilibrium for the complete and imperfect information $G^* = \{s_1, s_2, c_1, c_2, V_1, V_2, P^*\}$, which is analog to the game of incomplete information G, it is sufficient and necessary for the set of normalized strategies $s^* = (s_1^*, s_2^*)$ to be a Nash Equilibrium.

Holding the assumption of mutually consistent subjective probability distributions, $P_i^* = P^*(c), i = 1,2$, with know player type vectors $\{c_1, c_2\}$, a generalized bilateral bargaining model is elaborated subject to a set of all possible agreements defined as

$$X(d) = \{(x_1(c_1), x_2(c_2)) \,|\, x_1(c_1) + x_2(c_2) \leq z,\, x_i(c_i) \geq d_i\} \tag{2}$$

where $X \subset R^2$, $d_i \geq 0$, for all possible combinations of known types given by the Cartesian product $C_1 \times C_2$. Furthermore, given that the payoff for each player i, $x_i(c_i)$ is the weighted sum of all possible normalized strategies given c_i, the set of all possible agreements $X(d)$ is the weighted sum of all the possible $(k \times k)$ bargaining subsets of payoff vectors for all specific types, $c_1^l \in C_1$, $c_2^m \in C_2$, for all $l, m \in \{1,2,\ldots,k\}$:

$$X_{lm}(d) = \left\{ \left(x_1\left(c_1^l\right), x_2\left(c_2^m\right)\right) | \, x_1^l + x_2^m \leq z, x_i^\bullet \geq d_i^\bullet \right\}, \tag{3}$$

where each vector pair type specific bargaining subset $X_{lm}(d)$ adheres to Nash formulation and satisfies its conditions, such that there exists a rectangle T_{lm} formed by the separating hyperplane that bounds $X_{lm}(d)$ given a hyperbolic curve formed by the objective function of the bargaining solution $H = (x_1 - d_1)(x_2 - d_2)$. Thus, the product of all type pair specific bargaining sets, which is a convex combination of every set represented by (3), forms a set of all possible bargaining agreements such that for $C_1 \times C_2$

$$X(d) = \prod_{(l,m)\in(K,K)} X_{lm}(d), \quad K = \{1,2,\ldots,k\}$$

where the $\prod X_{lm}$ for $(l,m) \in (K,K)$ provides a weighted sum given the distribution of possible types for the common probability matrix P^*. Using Nash's axiom of

independence of irrelevant alternatives, which states that the optimal point $x^* \in X_{lm}$ is also the optimal point in T_{lm}, let the weighted space of all bounding T_{lm} rectangles that forms a convex hull in R_+^2 be

$$T = \prod_{(l,m)\in(K,K)} T_{lm}, \quad K = \{1,2,\ldots,k\}$$

satisfying convexity conditions necessary for existence and uniqueness of a Nash Equilibrium.

3 Unmeasurable Uncertainty and Inconsistency of Beliefs in Bilateral Bargaining

While Harsanyi's model with imperfect information holds under mutual consistency among players' probability distributions, Harsanyi [3] among others [5] recognized the possible failure of this model upon relaxing the assumption of mutual consistency. Specifically, Harsanyi acknowledge that by the nature of subjective probabilities even when two individuals have the same information and the same level of intelligence, thy may still assign different subjective probabilities to the same events. By relaxing the assumption of mutual consistency of the probability distributions among players, $P_1(c_1,c_2) \neq P_2(c_1,c_2)$. Thus, the ability to derive a common probability distribution P^* to construct an analogous game of complete but imperfect information G^* is no longer valid.

It is assumed that the type vectors, $c_i = \left(c_i^1, c_i^2, \ldots, c_i^k\right)$, are commonly known by both players while having different distributions of all the possible combinations of type vector pairs that span the space of the cartesian product $C_1 \times C_2$. Then for each player i according to their type vector, c_i, and the corresponding set of normalized strategies $s_i = s_i^*(c_i) = \left\{s_i^*\left(c_i^1\right), s_i^*\left(c_i^2\right), \ldots, s_i^*\left(c_i^k\right)\right\}$, we can define the game of incomplete information $G = \{S_1, S_2, C_1, C_2, V_1, V_2, P_1, P_2\}$ – where once again $V_i : [S_1 \times S_2 \times C_1 \times C_2] \rightarrow R$ represents the payoffs for player i in the normal for of the game.

Given the relaxation of mutual consistency of distributions, $P_1 \neq P^* \neq P_2$, each player i will assign different normalized payoffs to both players. The difference in normalized payoffs will produce a different weighted sum for each player. Thus, the normalized payoffs according to player i's distribution such that $x_i^i = V_i\left(s_1^{i*}(c_1), s_2^{i*}(c_2), c_1, c_2\right)$ and $x_j^i = V_j\left(s_1^{i*}(c_1), s_2^{i*}(c_2), c_1, c_2\right)$ for $i,j \in \{1,2\}, i \neq j$ will produce different agreement vectors than those produced according to player j's, .ie., $\left(x_1^i, x_2^i\right) \neq \left(x_1^j, x_2^j\right)$. Specifically, the expected payoffs in the normal form of the game will differ among players given that each player's weighted sum will assign a different probability weighting over the space of all possible combinations of type vector pairs, such that

$$E\left(x_i^i|\,c_i\right)=\int_{C_j}V_i\left(s_i^{i*},s_j^{i*},c_i,c_j\right)d_{(c_j)}P_i^i\left(c_j|c_i\right),E\left(x_j^i|c_j\right)=$$

$$=\int_{C_i}V_j\left(s_i^{i*},s_j^{i*},c_i,c_j\right)d_{(c_i)}P_j^i\left(c_i|c_j\right)$$

for $i,j\in\{1,2\},i\neq j$ where $P_j^i\left(c_i|c_j\right)$ is player i's estimate of the conditional probability player j assigns $\hat{P}_j=P_i^j$ for c_i given c_j. Thus, player i uses a sum expected payoffs which assigns a different probability weight for each possible normalized strategy, $E\left(x_j^j|c_j\right)\neq E\left(x_j^i|c_j\right)$.

Let us suppose that both players wrongly assume that their subjective probability distributions are mutually consistent such that $P_i^i=\hat{P}_i$ and $P_j^j=\hat{P}_j$ for $i,j\in\{1,2\},i\neq j$. Thus, each player i will be trying to solve a different game of imperfect information $G_i^*=\left\{S_1^{i*},S_2^{i*},C_1,C_2,V_1,V_2,P^{i*}\right\}$ where P^{i*} is what player i perceives to be the commonly known probability distribution over all possible type vectors. For the generalized bargaining model defined in the previous section with known type vectors $\{c_1,c_2\}$ and mutually inconsistent (unknown) probability distributions, $P_i\neq P_j\rightarrow P^{i*}\neq P^{j*}$, the set of all possible agreements given the set of expected payoffs for both players according to player i, $x^i=\left(x_i^i,x_j^i\right)$ such that

$$X^i(d)=\left\{\left(x_1^i\left(c_1\right),x_2^i\left(c_2\right)\right)|x_1^i\left(c_1\right)+x_2^i\left(c_2\right)\leq z,x_i^i\left(c_i\right)\geq d_i\right\}\tag{4}$$

for $d_i\geq 0,i=1,2$ will be different than the set of all possible agreements given the set of expected payoffs for both players according to player j, $x^j=\left(x_i^j,x_j^j\right)$ such that for $x^i\neq x^j\longrightarrow X^i(d)\neq X^j(d)$, where both are sets are in R^2.

While the set of all possible $(k\times k)$ type vector pair combinations from the cartesian product of all type vectors for each player, $C_1\times C_2$, remain the same, given different probability weighting of all normalized strategies, any particular combination of type vector pair, c_1^l,c_2^m for $l,m\in\{1,2,\ldots,k\}$ will produce different type vector specific bargaining sets, such that

$$X_{lm}^i(d)=\left\{\left(x_1^i\left(c_1^l\right),x_2^i\left(c_2^m\right)\right)|x_1^i\left(c_1^l\right)+x_2^i\left(c_2^m\right)\leq z,x_i^\bullet\geq d_i^\bullet\right\}\tag{5}$$

for $i=1,2;\ l,m\in\{1,2\}$, where (5) satisfies the conditions for Nash solution and it is bounded by a rectangle T_{lm}^i formed by a separating hyperplane produced by the hyperbolic curve formed by the objective function in Nash's bargaining solution, $H=\left(x_1^i-d_1\right)\left(x_2^i-d_2\right)$ according what player i perceives to be the vector of expected payoffs for both players, $x^i=\left(x_1^i,x_2^i\right)$. Thus, given different probability weighting for a specific vector type pair, $\left(c_1^l,c_2^m\right)$, $X_{lm}^i\neq X_{lm}^j$, their respective weighted sums of all possible type vector specific agreements will form different sets of all possible convex combinations in $C_1\times C_2$:

$$X^i(d)=\prod_{(l,m)\in(K,K)}X_{lm}^i(d)\neq X^j(d)=\prod_{(l,m)\in(K,K)}X_{lm}^j(d)$$

for $K = \{1,2,\ldots,k\}$. Thus, by Nash's axiom of Pareto optimality where $\left(x_1^{i*}\left(c_1^l\right), x_2^{i*}\left(c_2^m\right)\right) \in X_{lm}^i \rightarrow \left(x_1^{i*}\left(c_1^l\right), x_2^{i*}\left(c_2^m\right)\right) \in T_{lm}^i$, the weighted sum for each player i of all possible type vector specific agreements will produce a convex hull for the affine combination of all bounding rectangles $T_{lm}^i \in R_+^2$, where $T_{lm}^i \neq T_{lm}^j$ such that

$$T^i = \prod_{(l,m)\in(K,K)} T_{lm}^i \neq T^j = \prod_{(l,m)\in(K,K)} T_{lm}^j$$

for $i,j \in \{1,2\}, l,m \in \{1,2\}, i \neq j, l \neq m$. While both convex hulls formed by T^i and T^j satisfy the convexity conditions for the existence of an equilibrium, the expected equilibrium payoffs formed by both players will differ. Thus, leading to a bargaining process whereby neither Pareto nor Nash equilibria can be reached – aside from the disagreement point. That is, the weights assigned to their normalized strategies will lead to different expectations as to what the payoff demands should be according to each player's type. Thus, the affine combination of their demands could result in payoff demand vectors outside the space formed by the union of both spaces of all possible agreements. This breaks the convexity (preserving) condition necessary for the existence of an equilibrium, which leads to the inability of players reaching a bargaining solution.

3.1 Example of a Bilateral Bargaining Game with Incomplete Information and Inconsistency of Probability Distributions

Let us suppose two players are playing a game of dividing \$100 where the privately known payoff demand of player i is y_i^* for $i = 1,2$. If the sum of players' payoff demands do not exceed the given amount, i.e., $y_1^* + y_2^* \leq 100$, each player receives the payoff amount demanded, $y_i = y_i^*$. Otherwise, both players receive 0. Given the set of pure payoff demands and the set of types for each player, $c_i \in \{a_1, a_2\}$, the set of expected payoffs for player i are defined by

$$x_i = \frac{1}{2} y_i \times \mathbf{1}(c_i = a_1) + y_i \times \mathbf{1}(c_i = a_2)$$

where $\mathbf{1}(\cdot)$ is an indicator function and each player assigns a different probability to each possible type vector combination given by

$$C_1 \times C_2 = \{(a_1,a_1),(a_1,a_2),(a_2,a_1),(a_2,a_2)\}.$$

Let us assume that both players assign a negative correlation of types such that $Pr(c_j = a_m | c_i = a_l) = 1$ where $i,j \in \{1,2\}, l,m \in \{1,2\}, i \neq j, l \neq m$. If player 1 assigns an equal probability of having a pair of type vectors $(c_1,c_2) = (a_1,a_2)$ or $(c_1,c_2) = (a_2,a_1)$ such that $p_{12}^1 = p_{21}^1 = \frac{1}{2}$, then the vector of expected payoffs according to player 1 is given by

$$x^1 = \left(x_1^1, x_2^1\right) = \frac{1}{2} \cdot x^{1*} + \frac{1}{2} \cdot x^{1**}$$

where x^{1*} represents the possible payoff vectors given the pair of type vectors $(c_1,c_2)=(a_1,a_2)$ and x^{1**} represents the possible payoff vectors given the pair of type vectors $(c_1,c_2)=(a_2,a_1)$. Thus, player 1 will formulate a weighted sum of possible payoff vectors for each player such that

$$x_1^1=\frac{1}{2}\cdot\left(50-\frac{1}{2}x_2\right)+\frac{1}{2}(100-2x_2)$$
$$x_2^1=\frac{1}{2}\cdot(100-2x_1)+\frac{1}{2}\cdot\left(50-\frac{1}{2}x_1\right)$$

and form a space of all possible expected payoffs in the following manner.

- First, under the cooperative framework where player 1 receives the full amount $x_1=100$ while player 2 receives nothing $x_2=0$ when $(c_1,c_2)=(a_2,a_1)$ and player 1 receives nothing $x_1=0$ while player 2 receives the full amount $x_2=100$ with an equal probability of $\frac{1}{2}$, the expected Pareto optimal equilibrium payoff according to player 1 is $\left(x_1^1,x_2^1\right)=(50,50)$.
- Then, under the non-cooperative framework where each player's expected payoff is formed by the weighted sum of the Nash equilibrium outcomes under each pair of type vectors, the expected Nash equilibrium payoff according to player 1 is $\left(x_1^1,x_2^1\right)=(37.5,37.5)$.
- Finally, when players agree that player 1 receives the upper bound payment while player 2 receives the lowest bound payment the expected payoff vector according to player 1 is $\left(x_1^1,x_2^1\right)=(75,0)$. When players agree that player 1 receives the lowest bound payment while player 2 receives the upper bound payment the expect payoff according to player 1 is $\left(x_1^1,x_2^1\right)=(0,75)$.

Thus, the weighted space of all possible affine combinations of expected payoffs according to player 1 is represented by the convex hull shown in the figure below and is denominated as X^1.

Let us now assume that player 2 assigns unequal probabilities to the different pairs of type vectors where $(c_1,c_2)=(a_1,a_2)$ has a probability $p_{12}^2=\frac{3}{4}$ while $(c_1,c_2)=(a_2,a_1)$ has a probability $p_{21}^2=\frac{1}{4}$. The the vector of expected payoffs according to player 2 is given by

$$x^2=\left(x_1^2,x_2^2\right)=\frac{3}{4}\cdot x^{2*}+\frac{1}{4}\cdot x^{2**}$$

where x^{2*} represents the possible payoff vectors given the pair of type vectors $(c_1,c_2)=(a_1,a_2)$ and x^{2**} represents the possible payoff vectors given the pair of type vectors $(c_1,c_2)=(a_2,a_1)$. Thus, player 2 will formulate a weighted sum of possible payoff vectors for each player such that

$$x_1^2=\frac{3}{4}\cdot\left(50-\frac{1}{2}x_2\right)+\frac{1}{4}\cdot(100-2x_2)$$
$$x_2^2=\frac{3}{4}\cdot(100-2x_1)+\frac{1}{4}\cdot\left(50-\frac{1}{2}x_2\right)$$

and through a similar process but with different weights than player 1, the resulting space of all possible expected payoffs is formed by: (1) the expected Pareto optimal equilibrium payoff according to player 2, $\left(x_1^2, x_2^2\right) = (32.5, 60)$; (2) the expected Nash equilibrium according to player 2, $\left(x_1^2, x_2^2\right) = (31.25, 43.75)$; (3) the upper bound payment for player 1 and lower bound payment for player 2, $\left(x_1^2, x_2^2\right) = (62.5, 0)$, and the lower bound payment for player 1 and upper bound payment for player 2, $\left(x_1^2, x_2^2\right) = (0, 87.5)$. Thus, the weighted space of all possible affine combinations of expected payoffs according to player 2 is represented by the convex hull.

By having different probability distributions over the set of possible pairs of type vectors, $P_1 \neq P_2$, each player generates different weighted sums of possible outcomes. Thus, each player has different space of possible solutions with different probability weights assigned to their normalized strategies. Players will formulate different sets of optimal best-response strategies under the erroneous belief that both players have the same (mutually reinforcing) probability distribution. So even when both players have common knowledge of the structure of the game and agree to cooperate, given different distributions of probabilities of the different states, the models based on the assumption of rationality with mutual consistency of probability distributions fail to account for instance of nonexistence or uniqueness of an equilibrium.

4 Conclusions

As it was show in the previous example, we have a very different structure as compared to Nash and Harsanyi's models. The difference is explained by different distributions of the players' strategies. As it has been shown in experimental studies [1] [10] where players are privately induced with different distributions of all possible outcomes, when opposing players with different expectations of equilibrium payoff demands try to reach an agreement theory predictions are no longer consistent with observed behavior [1]. Far from attempting to discard existing models of bilateral bargaining of complete but imperfect information [2], we offer an alternative view of uncertainty to explain inconsistencies with theoretical predictions. Specifically, we develop a framework that accounts for varying degrees of knowledge and mutual consistency of beliefs in bilateral bargaining. Our goal in this particular analysis is to put forward a conjecture that establishes a direct relationship between the degree of inconsistency in beliefs among individuals and their ability to reach an agreement. Thus, providing the foundation for building a model of bilateral bargaining with incomplete information under an unmeasurable uncertainty framework [4].

Acknowledgements. This research was supported by the National Counsel of Science and Technology (CONACyT) of Mexico as part of the project SEP-2011 Ciencia Basica 169765; PROMEP 103.5/11/4330; PAICYT 464-10.

References

1. Cox, J.C., Shachat, J., Walker, M.: An experiment to evaluate Bayesian learning of Nash equilibrium plays. Games and Economic Behavior 34, 11–33 (2001)
2. Harsanyi, J.C.: Games with incomplete information played by 'Bayesian' players I-III. Management Sci. 14, 159–182, 320–334, 486–502 (1967)
3. Harsanyi, J.C., Selten, R.: A generalized Nash solution for two-person bargaining games with incomplete information. Management Sci. 18, 80–106 (1968)
4. Knight, F.H.: Risk, Uncertainty and Profit. The Riverside Press, Cambridge (1921)
5. Mertens, J.F., Zamir, S.: Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29 (1985)
6. Nash, J.F.: The bargaining problem. Econometrica 14, 155–162 (1950)
7. Osborne, M.J., Rubinstein, A.: Bargaining and Markets. Academic Press, New York (1990)
8. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press, Cambridge (1994)
9. Rubinstein, A.: "Afterword" of the 60th Anniversary Edition of J. Von Neumann O. Morgenstern's. In: Theory of Games and Economic Behavior, PDF File
10. Smith, V.L.: Bargaining and Market Behavior: Essays in Experimental Economics. Cambridge University Press, New York (2000)
11. Szidarovszky, F., Lin, C.H.: The alternating offer bargaining method under uncertainty. Applied Math. Comp. 76, 133–141 (1996)

An Interactive Fuzzy Satisficing Method for Multiobjective Stochastic Linear Programming Problems Considering Both Probability Maximization and Fractile Optimization

Hitoshi Yano

Abstract. In this paper, we propose an interactive fuzzy satisficing method for multiobjective stochastic linear programming problems to obtain a satisficing solution, in which the criteria of probability maximization and fractile optimization are considered simultaneously. In the proposed method, it is assumed that the decision maker has fuzzy goals for not only permissible objective levels of a probability maximization model but also permissible probability levels of a fractile criterion optimization model. After eliciting the corresponding membership functions for the fuzzy goals, two kinds of membership functions for permissible objective levels and permissible probability levels are integrated through the fuzzy decision. An interactive algorithm is proposed to obtain a satisficing solution from among a D_f-Pareto optimal solution set.

1 Introduction

In the real world decision making situations, we often have to make a decision under uncertainty. In order to deal with decision problems involving uncertainty, stochastic programming approaches have been developed [1],[2], and they were extended to multiobjective stochastic programming problems (MOSLP) [3],[6]. Recently, Sakawa et al. [4],[5] proposed interactive fuzzy satisficing methods for MOSLP to obtain a satisficing solution, in which a probability maximization model or a fractile optimization model are adopted in order to deal with the objective functions involving random variable coefficients. When adopting a probability maximization model or a fractile optimization model, it is required in advance that the decision maker specifies the parameters called permissible objective levels or permissible

Hitoshi Yano
Department of Social Sciences, School of Humanities and Social Sciences,
Nagoya City University, 1, Yamanohata, Mizuho-cho, Mizuho-ku, Nagoya, 467-8501 Japan
e-mail: yano@hum.nagoya-cu.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 119–128.
springerlink.com

probability levels. However, since there exist conflicts between permissible objective levels and permissible probability levels, it is required for the decision maker to specify appropriate values of permissible objective levels or permissible probability levels [7]. In this paper, we propose an interactive fuzzy satisficing method for MOSLP to obtain a satisficing solution, in which the criteria of probability maximization and fractile optimization are considered simultaneously. In section 2, we introduce D_f-Pareto optimal solution concept through the fuzzy decision [9], and formulate a minmax problem based on a fractile optimization model to obtain a D_f-Pareto optimal solution. In section 3, an interactive algorithm is proposed to obtain a satisficing solution from among D_f-Pareto optimal solution set. In section 4, the proposed algorithm is applied to a numerical example to demonstrate the feasibility of the proposed method under the hypothetical decision maker.

2 A Formulation of MOSLP through a Fractile Optimization Model

We focus on multiobjective stochastic linear programming problem (MOSLP) formally formulated as follows.

[MOSLP]

$$\min_{x \in X} \bar{z}(x) = (\bar{z}_1(x), \cdots, \bar{z}_k(x)) \tag{1}$$

where $x = (x_1, x_2, \cdots, x_n)^T$ is an n dimensional decision variable column vector whose elements $x_i, i = 1, \cdots, n$ are nonnegative, X is a linear constraint set with respect to x, $\bar{z}_i(x) = \bar{c}_i x + \bar{\alpha}_i, i = 1, \cdots, k$, are multiobjective functions involving random variable coefficients, $\bar{c}_i$ is a n dimensional random variable row vector expressed by $\bar{c}_i = c_i^1 + \bar{t}_i c_i^2$ where $\bar{t}_i$ is a random variable and $\bar{\alpha}_i$ is a random variable row vector expressed by $\bar{\alpha}_i = \alpha_i^1 + \bar{t}_i \alpha_i^2$. Let us denote distribution functions of random variables $\bar{t}_i$ as $T_i(\cdot)$. It is assumed that distribution functions $T_i(\cdot), i = 1, \cdots, k$ are strictly increasing and continuous.

If we adopt a fractile optimization model [4] for MOSLP, we can convert MOSLP to the following multiobjective programming problem, where $\hat{p} = (\hat{p}_1, \cdots, \hat{p}_k)$ is a k-dimensional vector of permissible probability levels which are specified by the decision maker, and $f_i, i = 1, \cdots, k$ are target variables for the objective functions.

[MOP1($\hat{p}$)]

$$\min_{x \in X, f_i \in R^1, i = 1, \cdots, k} (f_1, \cdots, f_k) \tag{2}$$

subject to

$$p_i(x, f_i) \geq \hat{p}_i, i = 1, \cdots, k \tag{3}$$

In MOP1($\hat{p}$), since a distribution function $T_i(\cdot)$ is continuous and strictly increasing, the constraints (3) can be transformed to the following form.

$$\begin{aligned} \hat{p}_i \leq p_i(x, f_i) &= T_i\left(\frac{f_i - (c_i^1 x + \alpha_i^1)}{c_i^2 x + \alpha_i^2}\right) \\ \Leftrightarrow \; f_i &\geq T_i^{-1}(\hat{p}_i) \cdot (c_i^2 x + \alpha_i^2) + (c_i^1 x + \alpha_i^1), i = 1, \cdots, k \end{aligned} \tag{4}$$

Let us define the right-hand side of the inequality (4) as follows.

$$f_i(x,\hat{p}_i) \stackrel{\text{def}}{=} T_i^{-1}(\hat{p}_i)\cdot(c_i^2x+\alpha_i^2)+(c_i^1x+\alpha_i^1) \tag{5}$$

Then, MOP1$(\hat{p})$ can be equivalently reduced to the following simple form.
[MOP2$(\hat{p})$**]**

$$\min_{x\in X}(f_1(x,\hat{p}_1),\cdots,f_k(x,\hat{p}_k)) \tag{6}$$

In order to deal with MOP2$(\hat{p})$, the decision maker must specify permissible probability levels $\hat{p}$ in advance. However, in general, the decision maker seems to prefer not only the less value of the objective function $f_i(x,\hat{p}_i)$ but also the larger value of the permissible probability level $\hat{p}_i$. From such a point of view, we consider the following multiobjective programming problem which can be regarded as a natural extension of MOP2$(\hat{p})$.
[MOP3]

$$\min_{x\in X,\hat{p}_i\in(0,1),i=1,\cdots,k}(f_1(x,\hat{p}_1),\cdots,f_k(x,\hat{p}_k),-\hat{p}_1,\cdots,-\hat{p}_k)$$

It should be noted in MOP3 that permissible probability levels $\hat{p}_i, i=1,\cdots,k$ are not fixed values but decision variables.

Considering the imprecise nature of the decision maker's judgment, we assume that the decision maker has a fuzzy goal for each objective function in MOP3. Such a fuzzy goal can be quantified by eliciting the corresponding membership function. Let us denote a membership function of an objective function $f_i(x,\hat{p}_i)$ as $\mu_{f_i}(f_i(x,\hat{p}_i))$, and a membership function of a permissible probability level $\hat{p}_i$ as $\mu_{\hat{p}_i}(\hat{p}_i)$ respectively. Then, MOP3 can be transformed to the following problem.
[MOP4]

$$\max_{x\in X,\hat{p}_i\in(0,1),i=1,\cdots,k}\big(\mu_{f_1}(f_1(x,\hat{p}_1)),\cdots,\mu_{f_k}(f_k(x,\hat{p}_k)),\mu_{\hat{p}_1}(\hat{p}_1),\cdots,\mu_{\hat{p}_k}(\hat{p}_k)\big)$$

Throughout this section, we make the following assumptions with respect to the membership functions $\mu_{f_i}(f_i(x,\hat{p}_i))$, $\mu_{\hat{p}_i}(\hat{p}_i), i=1,\cdots,k$.
Assumption 1.
$\mu_{\hat{p}_i}(\hat{p}_i), i=1,\cdots,k$ are strictly increasing and continuous with respect to $\hat{p}_i$, which is defined on the interval $[p_{i\min},p_{i\max}]\subset(0,1)\in\mathrm{R}^1$, where $\mu_{\hat{p}_i}(\hat{p}_i)=0$ if $0\le\hat{p}_i\le p_{i\min}$, and $\mu_{\hat{p}_i}(\hat{p}_i)=1$ if $p_{i\max}\le\hat{p}_i\le 1$.
Assumption 2.
$\mu_{f_i}(f_i(x,\hat{p}_i)), i=1,\cdots,k$ are strictly decreasing and continuous with respect to $f_i(x,\hat{p}_i)$, which is defined on the interval $[f_{i\min},f_{i\max}]\in\mathrm{R}^1$, where $\mu_{f_i}(f_i(x,\hat{p}_i))=0$ if $f_i(x,\hat{p}_i)\ge f_{i\max}$, and $\mu_{f_i}(f_i(x,\hat{p}_i))=1$ if $f_i(x,\hat{p}_i)\le f_{i\min}$.

In order to determine these membership functions appropriately, let us assume that the decision maker sets $p_{i\min}$, $p_{i\max}$ $f_{i\min}$ and $f_{i\max}$ as follows.

At first, the decision maker specifies $p_{i\min}$ and $p_{i\max}$ in his/her subjective manner, where $p_{i\min}$ is an acceptable minimum value and $p_{i\max}$ is a sufficiently satisfactory minimum value, and sets the intervals $P_i=[p_{i\min},p_{i\max}], i=1,\cdots,k$. Corresponding to the interval P_i, let us denote the interval of $\mu_{f_i}(f_i(x,\hat{p}_i))$ as $F_i(P_i)=[f_{i\min},f_{i\max}]$, where $f_{i\min}$ is a sufficiently satisfactory maximum value and $f_{i\max}$ is an acceptable

maximum value. $f_{i\min}$ can be obtained by solving the following linear programming problem.

$$f_{i\min} = \min_{x \in X} f_i(x, p_{i\min}) \tag{7}$$

In order to obtain $f_{i\max}$, we first solve the following k linear programming problems. $\min_{x \in X} f_i(x, p_{i\max}), i = 1, \cdots, k$. Let $x_i, i = 1, \cdots, k$ be the corresponding optimal solution. Using the optimal solutions $x_i, i = 1, \cdots, k$, $f_{i\max}$ can be obtained as follows [8].

$$f_{i\max} = \max_{\ell=1,\cdots,k,\ell \neq i} f_i(x_\ell, \hat{p}_{i\max}) \tag{8}$$

It should be noted here that, from (5), $\mu_{f_i}(f_i(x, \hat{p}_i))$ and $\mu_{\hat{p}_i}(\hat{p}_i)$ are conflict each other with respect to $\hat{p}_i$ for any $x \in X$. Here, let us assume that the decision maker adopts the fuzzy decision [9] in order to integrate both the membership functions $\mu_{f_i}(f_i(x, \hat{p}_i))$ and $\mu_{\hat{p}_i}(\hat{p}_i)$. Then, the integrated membership function $\mu_{D_{f_i}}(x, \hat{p}_i)$ can be defined as follows.

$$\mu_{D_{f_i}}(x, \hat{p}_i) = \min\{\mu_{\hat{p}_i}(\hat{p}_i), \mu_{f_i}(f_i(x, \hat{p}_i))\}, i = 1, \cdots, k \tag{9}$$

Using the membership functions $\mu_{D_{f_i}}(x, \hat{p}_i), i = 1, \cdots, k$, MOP4 can be transformed into the following form.

[MOP5]

$$\max_{x \in X, \hat{p}_i \in P_i, i=1,\cdots,k} \left(\mu_{D_{f_1}}(x, \hat{p}_1), \cdots, \mu_{D_{f_k}}(x, \hat{p}_k) \right)$$

In order to deal with MOP5, we introduce a D_f-Pareto optimal solution concept.

Definition 1.

$x^* \in X, \hat{p}_i^* \in P_i, i = 1, \cdots, k$ is said to be a D_f-Pareto optimal solution to MOP5, if and only if there does not exist another $x \in X, \hat{p}_i \in P_i, i = 1, \cdots, k$ such that $\mu_{D_{f_i}}(x, \hat{p}_i) \geq \mu_{D_{f_i}}(x^*, \hat{p}_i^*)$ $i = 1, \cdots, k$, with strict inequality holding for at least one i.

For generating a candidate of the satisficing solution which is also D_f-Pareto optimal, the decision maker is asked to specify the reference membership values [4],[5]. Once the reference membership values $\hat{\mu} = (\hat{\mu}_1, \cdots, \hat{\mu}_k)$ are specified, the corresponding D_f-Pareto optimal solution is obtained by solving the following min-max problem.

[MINMAX1($\hat{\mu}$)]

$$\min_{x \in X, \hat{p}_i \in P_i, i=1,\cdots,k, \lambda \in \mathrm{R}^1} \lambda \tag{10}$$

subject to

$$\hat{\mu}_i - \mu_{f_i}(f_i(x, \hat{p}_i)) \leq \lambda, i = 1, \cdots, k \tag{11}$$

$$\hat{\mu}_i - \mu_{\hat{p}_i}(\hat{p}_i) \leq \lambda, i = 1, \cdots, k \tag{12}$$

Because of Assumption 2 and $c_i^2 x + \alpha_i^2 > 0$, the constraints (11) can be transformed as follows.

$$\hat{p}_i \leq T_i \left(\frac{\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda) - (c_i^1 x + \alpha_i^1)}{c_i^2 x + \alpha_i^2} \right) \tag{13}$$

where $\mu_{f_i}^{-1}(\cdot)$ is an inverse function of $\mu_{f_i}(\cdot)$. From the constraints (12) and Assumption 1, it holds that $\hat{p}_i \geq \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda)$ where $\mu_{\hat{p}_i}^{-1}(\cdot)$ is an inverse function of $\mu_{\hat{p}_i}(\cdot)$. Therefore, the constraint (13) can be reduced to the following inequality where a permissible probability level $\hat{p}_i$ is disappeared.

$$\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda) - (c_i^1 x + \alpha_i^1) \geq T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda)) \cdot (c_i^2 x + \alpha_i^2) \tag{14}$$

Since the domain of $\mu_{f_i}^{-1}(\cdot)$ and $\mu_{\hat{p}_i}^{-1}(\cdot)$ is $[0,1] \in \mathrm{R}^1$, the feasible region of λ can be set as $\Lambda = [\max_{i=1,\cdots,k} \hat{\mu}_i - 1, \min_{i=1,\cdots,k} \hat{\mu}_i] \in \mathrm{R}^1$. Then, MINMAX1($\hat{\mu}$) can be equivalently reduced to the following problem.

[MINMAX2($\hat{\mu}$)]

$$\min_{x \in X, \lambda \in \Lambda} \lambda \tag{15}$$

subject to

$$\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda) - (c_i^1 x + \alpha_i^1) \geq T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda)) \cdot (c_i^2 x + \alpha_i^2), i = 1, \cdots, k \tag{16}$$

It should be noted here that the constraints (16) can be reduced to a set of linear inequalities for some fixed value $\lambda \in \Lambda$. This means that an optimal solution (x^*, λ^*) of MINMAX2($\hat{\mu}$) is obtained by combined use of the bisection method with respect to λ and the first-phase of the two-phase simplex method of linear programming.

The relationship between the optimal solution (x^*, λ^*) of MINMAX2($\hat{\mu}$) and D_f-Pareto optimal solutions can be characterized by the following theorem.

Theorem 1.

(1) If $x^* \in X, \lambda^* \in \Lambda$ is a unique optimal solution of MINMAX2($\hat{\mu}$), then $x^* \in X, \hat{p}_i^* = \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*) \in P_i, i = 1, \cdots, k$ is a D_f-Pareto optimal solution.

(2) If $x^* \in X, \hat{p}_i^* \in P_i, i = 1, \cdots, k$ is a D_f-Pareto optimal solution, then $x^* \in X$, $\lambda^* = \hat{\mu}_i - \mu_{\hat{p}_i}(\hat{p}_i^*) = \hat{\mu}_i - \mu_{f_i}(f_i(x^*, \hat{p}_i^*)), i = 1, \cdots, k$ is an optimal solution of MINMAX2($\hat{\mu}$) for some reference membership values $\hat{\mu} = (\hat{\mu}_1, \cdots, \hat{\mu}_k)$.

(Proof)

(1) From (16), it holds that $\hat{\mu}_i - \lambda^* \leq \mu_{f_i}(f_i(x^*, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*))), i = 1, \cdots, k$. Assume that $x^* \in X, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*) \in P_i, i = 1, \cdots, k$ is not a D_f-Pareto optimal solution. Then, there exist $x \in X, \hat{p}_i \in P_i, i = 1, \cdots, k$ such that $\mu_{D_{f_i}}(x, \hat{p}_i) = \min\{\mu_{\hat{p}_i}(\hat{p}_i), \mu_{f_i}(f_i(x, \hat{p}_i)), \} \geq \mu_{D_{f_i}}(x^*, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) = \hat{\mu}_i - \lambda^*, i = 1, \cdots, k$, strict inequality holding for at least one i. Then it holds that

$$\mu_{\hat{p}_i}(\hat{p}_i) \geq \hat{\mu}_i - \lambda^*, i = 1, \cdots, k, \tag{17}$$

$$\mu_{f_i}(f_i(x, \hat{p}_i)) \geq \hat{\mu}_i - \lambda^*, i = 1, \cdots, k. \tag{18}$$

From the definition (5), the inequalities (17) and (18) can be transformed into the inequalities, $\hat{p}_i \geq \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*), \hat{p}_i \leq T_i\left(\dfrac{\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda^*) - (c_i^1 x^* + \alpha_i^1)}{c_i^2 x^* + \alpha_i^2}\right)$. This means

that there exists some $x \in X$ such that $\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda^*) - (c_i^1 x + \alpha_i^1) \geq T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x + \alpha_i^2), i = 1, \cdots, k$, which contradicts the fact that $x^* \in X, \lambda^* \in \Lambda$ is a unique optimal solution to MINMAX2($\hat{\mu}$).
(2) Assume that $x^* \in X, \lambda^* \in \Lambda$ is not an optimal solution to MINMAX2($\hat{\mu}$) for any reference membership values $\hat{\mu} = (\hat{\mu}_1, \cdots, \hat{\mu}_k)$, which satisfies the inequalities $\hat{\mu}_i - \lambda^* = \mu_{\hat{p}_i}(\hat{p}_i^*) = \mu_{f_i}(f_i(x^*, \hat{p}_i^*)), i = 1, \cdots, k$. Then, there exists some $x \in X, \lambda < \lambda^*$ such that

$$\begin{aligned} &\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda) - (c_i^1 x + \alpha_i^1) \geq T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda)) \cdot (c_i^2 x + \alpha_i^2), \\ \Leftrightarrow\ &\mu_{f_i}(f_i(x, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda)) \geq \hat{\mu}_i - \lambda, i = 1, \cdots, k. \end{aligned} \tag{19}$$

Because of (19) and $\hat{\mu}_i - \lambda > \hat{\mu}_i - \lambda^*, i = 1, \cdots, k$, the following inequalities hold.

$$\begin{aligned} \mu_{\hat{p}_i}(\mu_{\hat{p}_i}^{-1}(\mu_i - \lambda)) &> \mu_{\hat{p}_i}(\hat{p}_i^*), i = 1, \cdots, k, \\ \mu_{f_i}(f_i(x, \mu_{\hat{p}_i}^{-1}(\mu_i - \lambda))) &> \mu_{f_i}(f_i(x^*, \hat{p}_i^*)), i = 1, \cdots, k. \end{aligned}$$

This contradicts the fact that $x^* \in X, \hat{p}_i^* \in P_i, i = 1, \cdots, k$ is a D_f-Pareto optimal solution.

By adopting a probability maximization model [5] for MOSLP, a Pareto optimality concept such as a D_f-Pareto optimal solution can be defined, and the similar problem such as MINMAX2($\hat{\mu}$) to obtain a Pareto optimal solution can be formulated.

3 An Interactive Algorithm

In this section, we propose an interactive algorithm to obtain a satisficing solution from among a D_f-Pareto optimal solution set. Unfortunately, it is not guaranteed that the optimal solution (x^*, λ^*) of MINMAX2($\hat{\mu}$) is D_f-Pareto optimal, if (x^*, λ^*) is not unique. In order to guarantee the D_f-Pareto optimality, we first assume that k constraints (16) of MINMAX2($\hat{\mu}$) are active at the optimal solution (x^*, λ^*), *i.e.*,

$$\mu_{f_i}^{-1}(\hat{\mu}_i - \lambda^*) - (c_i^1 x^* + \alpha_i^1) = T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x^* + \alpha_i^2), i = 1, \cdots, k. \tag{20}$$

If the j-th constraint of (16) is inactive, *i.e.*,

$$\begin{aligned} &\mu_{f_j}^{-1}(\hat{\mu}_j - \lambda^*) - (c_j^1 x^* + \alpha_j^1) > T_j^{-1}(\mu_{\hat{p}_j}^{-1}(\hat{\mu}_j - \lambda^*)) \cdot (c_j^2 x^* + \alpha_j^2), \\ \Leftrightarrow\ &\mu_{f_j}^{-1}(\hat{\mu}_j - \lambda^*) > f_j(x^*, \mu_{\hat{p}_j}^{-1}(\hat{\mu}_j - \lambda^*)), \end{aligned} \tag{21}$$

we can convert the inactive constraint (21) into the active one by applying the bisection method, where $G_j(\hat{\mu}_j) \stackrel{\text{def}}{=} \mu_{f_j}^{-1}(\hat{\mu}_j - \lambda^*) - f_j(x^*, \mu_{\hat{p}_j}^{-1}(\hat{\mu}_j - \lambda^*))$.

[The bisection method for the inactive constraint]
Step 1. Set $q_j^L \leftarrow \hat{\mu}_j, q_j^R \leftarrow \lambda^* + 1$.
Step 2. Set $q_j \leftarrow (q_j^L + q_j^R)/2$.
Step 3. If $G_j(q_j) > 0$ then $q_j^L \leftarrow q_j$ and go to Step 2, else if $G_j(q_j) < 0$ then $q_j^R \leftarrow q_j$ and go to Step 2, else update the reference membership value as $\hat{\mu}_j \leftarrow q_j$ and stop.

For the optimal solution (x^*, λ^*) of MINMAX2($\hat{\mu}$), where the active conditions (20) are satisfied, we solve the D_f-Pareto optimality test problem defined as follows.
[Test problem for D_f-Pareto optimality]

$$\max_{x \in X, \varepsilon_i \geq 0, i=1,\cdots,k} w = \sum_{i=1}^{k} \varepsilon_i \tag{22}$$

subject to

$$\begin{aligned} & T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x + \alpha_i^2) + (c_i^1 x + \alpha_i^1) + \varepsilon_i \\ = \; & T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x^* + \alpha_i^2) + (c_i^1 x^* + \alpha_i^1), i = 1, \cdots, k \end{aligned} \tag{23}$$

For the optimal solution of the above test problem, the following theorem holds.

Theorem 2.
For the optimal solution $\check{x} \in X, \check{\varepsilon}_i \geq 0, i = 1, \cdots, k$ of the test problem (22)-(23), if $w = 0$ (equivalently, $\check{\varepsilon}_i = 0, i = 1, \cdots, k$), then $x^* \in X, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*) \in P_i, i = 1, \cdots, k$ is a D_f-Pareto optimal solution.
(Proof)
From the active conditions (20), it holds that $\hat{\mu}_i - \lambda^* = \mu_{f_i}(f_i(x^*, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*))) = \mu_{\hat{p}_i}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)), i = 1, \cdots, k$. If $x^* \in X, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*) \in P_i, i = 1, \cdots, k$ is not a D_f-Pareto optimal solution, there exists some $x \in X, \hat{p}_i \in P_i, i = 1, \cdots, k$ such that $\mu_{D_{f_i}}(x, \hat{p}_i) = \min\{\mu_{\hat{p}_i}(\hat{p}_i), \mu_{f_i}(f_i(x, \hat{p}_i))\} \geq \mu_{D_{f_i}}(x^*, \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) = \hat{\mu}_i - \lambda^*, i = 1, \cdots, k$, with strict inequality holding for at least one i. This means that the following inequalities hold.

$$\mu_{\hat{p}_i}(\hat{p}_i) \geq \hat{\mu}_i - \lambda^*, i = 1, \cdots, k \tag{24}$$

$$\mu_{f_i}(f_i(x, \hat{p}_i)) \geq \hat{\mu}_i - \lambda^*, i = 1, \cdots, k \tag{25}$$

This means that there exists some $x \in X$ such that

$$\begin{aligned} & \mu_{f_i}^{-1}(\hat{\mu}_i - \lambda^*) \geq (c_i^1 x + \alpha_i^1) + T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x + \alpha_i^2), \\ \Leftrightarrow \; & T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x^* + \alpha_i^2) + (c_i^1 x^* + \alpha_i^1) \\ & \geq T_i^{-1}(\mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)) \cdot (c_i^2 x + \alpha_i^2) + (c_i^1 x + \alpha_i^1), i = 1, \cdots, k, \end{aligned} \tag{26}$$

with strict inequality holding for at least one i. This contradicts the fact that $w = 0$.

Now, following the above discussions, we can present the interactive algorithm in order to derive a satisficing solution from among a D_f-Pareto optimal solution set.

[An interactive algorithm]
Step 1: The decision maker specifies $p_{i\min}$ and $p_{i\max}, i = 1, \cdots, k$ in his/her subjective manner. On the interval $P_i = [p_{i\min}, p_{i\max}]$, the decision maker sets his/her membership functions $\mu_{\hat{p}_i}(\hat{p}_i), i = 1, \cdots, k$ according to Assumption 1.
Step 2: Corresponding to the interval P_i, compute $f_{i\min}$ and $f_{i\max}$ by solving the problems (7) and (8). On the interval $F_i(P_i) = [f_{i\min}, f_{i\max}]$, the decision maker sets his/her membership functions $\mu_{f_i}(f_i(x, \hat{p}_i)), i = 1, \cdots, k$ according to Assumption 2.
Step 3: Set the initial reference membership values as $\hat{\mu}_i = 1, i = 1, \cdots, k$.
Step 4: Solve MINMAX2($\hat{\mu}$) by combined use of the bisection method and the first-phase of the two-phase simplex method of linear programming. For the optimal solution (x^*, λ^*), D_f-Pareto optimality test problem is solved.
Step 5: If the decision maker is satisfied with the current values of the D_f-Pareto optimal solution $\mu_{D_{f_i}}(x^*, \hat{p}_i^*), i = 1, \cdots, k$, where $\hat{p}_i^* = \mu_{\hat{p}_i}^{-1}(\hat{\mu}_i - \lambda^*)$, then stop. Otherwise, the decision maker must update his/her reference membership values $\hat{\mu}_i, i = 1, \cdots, k$, and return to Step 4.

4 A Numerical Example

We consider the following three-objective stochastic linear programming problems to demonstrate the feasibility of the proposed method, which is formulated by Sakawa et al. [5],
[MOSLP]

$$\min\ (c_1^1 + \bar{t}_1 c_1^2)x + (\alpha_1^1 + \bar{t}_1 \alpha_1^2)$$
$$\min\ (c_2^1 + \bar{t}_2 c_2^2)x + (\alpha_2^1 + \bar{t}_2 \alpha_2^2)$$
$$\min\ (c_3^1 + \bar{t}_3 c_3^2)x + (\alpha_3^1 + \bar{t}_3 \alpha_3^2)$$

subject to $x \in X \stackrel{\text{def}}{=} \{x \geq 0 \mid a_i x \leq b_i, i = 1, \cdots, 7\}$

where $x = (x_1, x_2, \cdots, x_{10})^T$ is a ten-dimensional decision column vector. The values of $a_i, i = 1, \cdots, 7, c_1^1, c_1^2, c_2^1, c_2^2, c_3^1, c_3^2, \alpha_1^1, \alpha_1^2, \alpha_2^1, \alpha_2^2, \alpha_3^1, \alpha_3^2$ are shown in Table 1. $b_1 = 132.907, b_2 = -222.897, b_3 = -196.624, b_4 = 70.8059, b_5 = -167.619, b_6 = 124.543, b_7 = 88.1748$. $\bar{t}_i, i = 1, 2, 3$ are Gaussian random variables defined as $\bar{t}_1 \sim \mathrm{N}(4, 2^2), \bar{t}_2 \sim \mathrm{N}(3, 3^2), \bar{t}_3 \sim \mathrm{N}(3, 2^2)$.

Let us assume that the hypothetical decision maker sets $P_i = [p_{i\min}, p_{i\max}] = [0.5, 0.9]$, $i = 1, 2, 3$, and sets the linear membership functions $\mu_{\hat{p}_i}(\hat{p}_i) = \dfrac{\hat{p}_i - p_{i\min}}{p_{i\max} - p_{i\min}}$, $i = 1, 2, 3$ for the permissible probability levels (Step 1). For the intervals P_i, $F_i(P_i) = [f_{i\min}, f_{i\max}], i = 1, 2, 3$ are obtained by solving the problems (7) and (8) as $[f_{1\min}, f_{1\max}] = [1855.57, 2599.30]$, $[f_{2\min}, f_{2\max}] = [340.617, 1066.26]$, $[f_{3\min}, f_{3\max}] = [-1067.25, -610.939]$ respectively. On the intervals $F_i(P_i), i = 1, 2, 3$, the hypothetical decision maker sets the linear membership functions $\mu_{f_i}(f_i(x, \hat{p}_i)) = \dfrac{f_{i\max} - f_i(x, \hat{p}_i)}{f_{i\max} - f_{i\min}}, i = 1, 2, 3$ (Step 2). Set the initial reference membership values as $\hat{\mu}_i = 1, i = 1, \cdots, k$ (Step 3), and solve

Table 1 Constant coefficients of objective functions and constraints

x	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
c_1^1	19	48	21	10	18	35	46	11	24	33
c_1^2	3	2	2	1	4	3	1	2	4	2
c_2^1	12	-46	-23	-38	-33	-48	12	8	19	20
c_2^2	1	2	4	2	2	1	2	1	2	1
c_3^1	-18	-26	-22	-28	-15	-29	-10	-19	-17	-28
c_3^2	2	1	3	2	1	2	3	3	2	1
a_1	12	-2	4	-7	13	-1	-6	6	11	-8
a_2	-2	5	3	16	6	-12	12	4	-7	-10
a_3	3	-16	-4	-8	-8	2	-12	-12	4	-3
a_4	-11	6	-5	9	-1	8	-4	6	-9	6
a_5	-4	7	-6	-5	13	6	-2	-5	14	-6
a_6	5	-3	14	-3	-9	-7	4	-4	-5	9
a_7	-3	-4	-6	9	6	18	11	-9	-4	7

Table 2 Interactive processes

@	iteration 1	iteration 2	iteration 3
$\hat{\mu}_1$	1	0.61	0.59
$\hat{\mu}_2$	1	0.54	0.525
$\hat{\mu}_3$	1	0.56	0.535
$\mu_{D_{f_1}}(x^*, \hat{p}_1^*)$	0.560356	0.591202	0.594255
$\mu_{D_{f_2}}(x^*, \hat{p}_2^*)$	0.560356	0.521202	0.529255
$\mu_{D_{f_3}}(x^*, \hat{p}_3^*)$	0.560356	0.541202	0.539255
$\hat{p}_1^*$	0.724142	0.736481	0.737702
$\hat{p}_2^*$	0.724142	0.708481	0.711702
$\hat{p}_3^*$	0.724142	0.716481	0.715702
$f_1(x^*, \hat{p}_1^*)$	2182.55	2159.61	2157.34
$f_2(x^*, \hat{p}_2^*)$	659.642	688.054	682.211
$f_3(x^*, \hat{p}_3^*)$	-866.634	-857.894	-857.005

MINMAX2$(\hat{\mu})$ to obtain the optimal solution (x^*, λ^*) (Step 4). The interactive processes of the hypothetical decision maker are shown in Table 2. At the first iteration, the optimal solution $\mu_{D_{f_i}}(x^*, \hat{p}_i^*) = 0.560356, i = 1,2,3$ is obtained. In order to improve $\mu_{D_{f_1}}(x^*, \hat{p}_1^*)$ at the expense of $\mu_{D_{f_2}}(x^*, \hat{p}_2^*)$, the hypothetical decision maker updates the reference membership values as $(\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3) = (0.61, 0.54, 0.56)$. In this example, at the third iteration, a satisficing solution of the decision maker is obtained.

5 Conclusion

In this paper, an interactive fuzzy satisficing method for MOSLP is proposed to obtain a satisficing solution, in which the criteria of probability maximization and fractile optimization are considered simultaneously. In the proposed method, on the basis of a fractile optimization model, D_f-Pareto optimal solution concept is introduced, which is obtained by solving MINMAX2($\hat{\mu}$). By a similar way, on the basis of a probability maximization model, we can define a Pareto optimality concept such as D_f-Pareto optimality and formulate minmax problem such as MINMAX2($\hat{\mu}$). In our proposed method, it is expected to obtain a satisficing solution, in which not only the proper balance between permissible objective values and permissible probability levels are attained through the fuzzy decision, but also the preference of the decision maker for the objective functions in MOSLP are reflected through the reference membership values.

References

1. Charnes, A., Cooper, W.W.: Chance Constrained Programming. Management Science 6, 73–79 (1959)
2. Danzig, G.B.: Linear Programming under Uncertainty. Management Science 1, 197–206 (1955)
3. Stancu-Minasian, I.M.: Overview of Different Approaches for Solving Stochastic Programming Problems with Multiple Objective Functions. In: Slowinski, R., Teghem, J. (eds.) Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming under Uncertainty, pp. 71–101. Kluwer Academic Publishers (1990)
4. Kato, K., Katagiri, H., Sakawa, M.: An Interactive Fuzzy Satisficing Method for Multiobjective Stochastic Programming Problems Using Fractile Optimization Model. In: The 10th IEEE International Conference on Fuzzy Systems 3, Melbourne, pp. 25–31 (2001)
5. Sakawa, M., Kato, K., Katagiri, H.: An Interactive Fuzzy Satisficing Method for Multiobjective Linear Programming Problems with Random Variable Coefficients Through a Probability Maximization Model. Fuzzy Sets and Systems 146, 205–220 (2004)
6. Teghem, J., Dufrane, D., Thauvoye, M.: STRANGE: An Interactive Method for Multi-Objective Linear Programming Under Uncertainty. European Journal of Operational Research 26, 65–82 (1986)
7. Yano, H.: Fuzzy Approaches for Multiobjective Stochastic Linear Programming Problems Considering Both Probability Maximization and Fractile Optimization. In: Proceedings of 2011 IEEE International Conference on Fuzzy Systems, vol. 23, pp. 1866–1873 (2011)
8. Zimmermann, H.-J.: Fuzzy Programming and Linear Programming with Several Objective Functions. Fuzzy Sets and Systems 1, 45–55 (1978)
9. Zimmermann, H.-J.: Fuzzy Sets, Decision-Making and Expert Systems. Kluwer Academic Publishers (1987)

Assignment Problem Based on Mathematical Formation of Consensus

Hiroaki Ishii and Yung-Lung Lee

Abstract. In this paper, we extend the distance measure of Cook and Kress and propose a procedure to choose a suitable person to each suitable position. When we choose the most suitable person among a number of them, we often use voting with preference of persons to positions. First using pairwise comparison, importance values of positions are calculated. After that, based on preference data of voters, we make a consensus formation extending distance measure of Cook and Kress. That is, the total sum of weighted distance between each voter's preference of positions and the final preference of them to positions are to be minimized. This problem is formulated as a special transportation problem and this consensus formation is found using some solution method for the transportation problem. Next we discuss the applicability to selection of players to positions for some sports. Finally we summarize the paper and discuss further research problem.

Keywords: Assignment of candidates to positions, AHP, Voting data, Distance measure, Assignment problem, Consensus making.

1 Introduction

Rank ordering is used to aggregate some preference of voters and choose the most desirable candidates. It is applied to marketing and policy assessment as well as polling and many unique researches are done recently such as using linguistic data [Herrera(1996)]. Recently sport is promising area for its application. In this paper we propose a mathematical selection method for candidates applicable to sports.

Hiroaki Ishii
School of Science and Technology, Kwansei Gakuin University2-1
Gakuen Sanda Hyogo 669-1337, Japan

Yung-Lung Lee
Department of Land Management and Development, Chang Jung Christian University
396-1 Changong Road, Gueiren Township, Tainan County 71101,Taiwan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 129–134.
springerlink.com

Voting systems have been widely studied in group decision making. The social choice theory is considered to be the first theory that tackled through [Borda (1781)] and [Condorcet (1785)] works. There are many approaches on mathematical approach such as Data Envelope Analysis (DEA) or Analytic Hierarchy Process (AHP) or distance measure.

It is important to consider the meaning of "winner" since it is different between voting systems. The attitude of superior winning is based on majority rule of Condorcet. The winner is decided according to the number of votes the candidate obtains, that is, the number is large or small and the way to win by wide or narrow margin is unconcerned. Borda's theory is based on the attitude to select the highly-positioned candidate by many voters and globally balanced winner is chosen. While each candidate has weights applied to her/his standings in Borda's theory, she/he is permitted to choose the most favorable weights in DEA manner.

The distance measure is based on a different idea that individual preference for a set of candidates should be aggregated (see related works [Cook(2006)], [Cook and Kress(1984)], [Green et al.(1996)]). Tanabe and Ishii [Tanabe(2007)] extended the distance measure to construct a joint ballot model. We further extend this model and consider the position in place of ranking and preference of experts. Section 2 formulates some assignment problem of candidates to positions. Section 3 proposes how to tackle this problem and finally its optimal assignment is found by some linear assignment problem. Section 4 summarize this paper and discusses further research problem.

2 Formulation of Our Problem

We consider the case that there are r positions to be assigned from n candidates. To each position only one candidate is assigned and it is determined by the voting data of m experts, that is, each expert has a preference of candidate about each position. This preference of the expert ℓ is denoted by the $n \times r$ preference matrix

$$A^{\ell} = (\alpha_{ij}^{\ell}),\ \ell = 1, 2, ..., m$$

where

$$\alpha_{ij}^{\ell} = \begin{cases} 1 & \text{if the expert } \ell \text{ consider candidate } i \text{ is suitable to position } j \\ 0 & \text{otherwise} \end{cases}$$

Note that usually $r \le n$ different from the model of Cook[Cook, 2006]. Distance between preference matrices $A = (a_{ij})$ and $B = (b_{ij})$ is defined as $\sum_{i=1}^{n} \sum_{j=1}^{r} | a_{ij} - b_{ij} |$. We consider the aggregated decision among experts is denoted by the preference matrix $X = (x_{ij})$ and it should be determined by

minimizing the total weighted distance defined as $\sum_{\ell=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{r} w_j \mid \alpha_{ij}^{\ell} - x_{ij} \mid$ where w_j is the importance of the position *j*. w_j is determined from the pair-wise comparison between each pair of positions by using AHP as follows:

First construct $r \times r$ pair-wise comparison matrix $D = (d_{ij})$ with respect to positions where

$$d_{ij} = \begin{cases} 1 & (i = j) \\ \dfrac{1}{d_{ji}} & (i \neq j) \end{cases}$$

and d_{ij} denotes how many times important position *i* to position *j*, using the number among the number 1to9. Here we consider ideally $d_{ij} = \dfrac{w_i}{w_j}$ as a ratio between weights w_i and w_j.

Therefore *j* th element of an eigen vector corresponding to eigen value *r* with respect to matrix *C* gives weight w_j since

$$\begin{pmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_j} & \cdots & \frac{w_1}{w_r} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_j} & \cdots & \frac{w_2}{w_r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{w_i}{w_1} & \cdots & \frac{w_i}{w_i} & \cdots & \frac{w_i}{w_j} & \cdots & \frac{w_i}{w_r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{w_r}{w_1} & \cdots & \cdots & \cdots & \frac{w_r}{w_j} & \cdots & \frac{w_r}{w_r} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_j \\ \vdots \\ w_r \end{pmatrix} = r \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_j \\ \vdots \\ w_r \end{pmatrix}$$

3 Optimal Decision of Assignment

Since the element of preference matrix is 0 or 1, minimizing $\sum_{\ell=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{r} w_j \mid \alpha_{ij}^{\ell} - x_{ij} \mid$ is equivalent to maximizing $\sum_{\ell=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{r} w_j \alpha_{ij}^{\ell} \cdot x_{ij}$ as is easily shown. Cook and Kress [Cook and Kress (1984)] has constructed relative

positioning to express the degree of differences and this is basically equivalent to the difference $\sum_{\ell=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{r}|\alpha_{ij}^{\ell}-x_{ij}|$ before weighting.

The position j forward indicator vector $P^{+}(j)$ and the position j backward indicator vector $P^{-}(j)$ are

$$P^{+}(j)=\left[\sum_{t=j+1}^{r}\alpha_{it}\right],\ P^{-}(j)=\left[\sum_{t=1}^{j-1}\alpha_{it}\right] \tag{1}$$

For an example of preference matrix A given as below, candidate 1 is suitable to position 2, B position 3, C position 1 and D position 4, but E is not suitable to any position

Table 1 An example of preference matrix for 5 candidates, A,B,C,D,E and 4 positions

	Position			
	1	2	3	4
A	0	1	0	0
B	0	0	1	0
Candidate C	1	0	0	0
D	0	0	0	1
E	0	0	0	0

Then ,

$$P^{+}(1)=\begin{bmatrix}1\\1\\0\\1\\0\end{bmatrix},\quad P^{+}(2)=\begin{bmatrix}0\\1\\0\\1\\0\end{bmatrix},\ P^{+}(3)=\begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix},\quad P^{+}(4)=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix},$$

$$P^{-}(1)=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix},\ P^{-}(2)=\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}\ P^{-}(3)=\begin{bmatrix}1\\0\\1\\0\\0\end{bmatrix},\quad P^{-}(4)=\begin{bmatrix}1\\1\\1\\0\\0\end{bmatrix}$$

The relative distance function $d_P(A,B)$ is given with $P_A^{+}(j),P_B^{+}(j)$ and $P_A^{-}(j),P_B^{-}(j)$ as below.

$$d_P(A,B)=r(r-1)-\sum_{J=1}^{r}\left[\left\langle P_A^{+}(j),P_B^{+}(j)\right\rangle+\left\langle P_A^{-}(j),P_B^{-}(j)\right\rangle\right]\quad .$$

where $\left\langle P_A^+(j),\ P_B^+(j)\right\rangle$ is an inner product between $P_A^+(j)$ and $P_B^+(j)$, and $\left\langle P_A^-(j), P_B^-(j)\right\rangle$ that between $P_A^-(j)$ and $P_B^-(j)$. Minimizing the relative distance is equivalent to maximizing the sum of an inner product $\sum_{j=1}^{r}\left[\left\langle P_A^+(j), P_B^+(j)\right\rangle+\left\langle P_A^-(j), P_B^-(j)\right\rangle\right]$. So we seek the assignment matrix (preference matrix) $X=(x_{ij})$ maximizing the total weighted sum $\sum_{\ell=1}^{m}\sum_{j=1}^{r} w_j\left[\left\langle P_{A^\ell}^+(j), P_X^+(j)\right\rangle+\left\langle P_{A^\ell}^-(j), P_X^-(j)\right\rangle\right]$. Rewriting it and divided by (*r*-1) results $\sum_{\ell=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{r} w_j\left[(\sum_{t=j+1}^{r}\alpha_{it}^{\ell})(\sum_{t=j+1}^{r}x_{it})+(\sum_{t=1}^{j-1}\alpha_{it}^{\ell})(\sum_{t=1}^{j-1}x_{it})\right]$ and further by changing the order of summation, finally we have

$$\sum_{i=1}^{n}\sum_{k=1}^{r}(\sum_{\ell=1}^{m} w_k\{\sum_{j=1}^{k-1}(\sum_{t=j+1}^{r}\alpha_{it}^{\ell})+\sum_{j=k+1}^{r}(\sum_{t=1}^{j-1}\alpha_{it}^{\ell})\})x_{ik} \tag{2}$$

Then setting $c_{ik}=w_k\{\sum_{j=1}^{k-1}(\sum_{t=j+1}^{r}\alpha_{it}^{\ell})+\sum_{j=k+1}^{r}(\sum_{t=1}^{j-1}\alpha_{it}^{\ell})\}, i=1,2,...,n, k=1,...,r$, we have the following problem to be solved.

$$\begin{aligned}&\text{Maximize } \sum_{i=1}^{n}\sum_{k=1}^{r} c_{ik}\,x_{ik}\\ &\text{subject to } \sum_{i=1}^{n} x_{ik}=1,\ k=1,2,...,r,\ \sum_{k=1}^{r} x_{ik}\le 1,\ i=1,2,...,n,\\ &x_{ik}=0\,or\,1,\ i=1,2,...,n, k=1,2,..,r\end{aligned} \tag{3}$$

Further setting $C=\sum_{i=1}^{n}\sum_{k=1}^{r} c_{ik}$ and

$$c'_{ik}=\begin{cases} C-c_{ik}, & k=1,2,...,r\\ M\ (big\ number),\ k=r+1,...,n\end{cases}\quad i=1,2,...,n,$$

Problem (3) is further transformed into the following linear assignment problem.

$$\begin{aligned}&\text{Minimize } \sum_{i=1}^{n}\sum_{k=1}^{n} c'_{ik}\,x_{ik}\\ &\text{subject to } \sum_{i=1}^{n} x_{ik}=1,\ k=1,2,...,n,\ \sum_{k=1}^{n} x_{ik}=1,\ i=1,2,...,n,\\ &x_{ik}=0\,or\,1,\ i=1,2,...,n, k=1,2,..,n\end{aligned} \tag{4}$$

Problem (4) can be solved efficiently using some algorithm for an assignment problem such as [Munkres (1957)] (Also refer to [Pentico(2007)]. From an optimal solution, we obtain an effective assignment of candidates to positions. Note that we can solve our problem as a transportation problem such as problem (3) in place of transformation to (4) though there needs a small wisdom.

4 Conclusions and Future Remarks

We have considered an extended model of Tanabe[Tanabe (2007)] with not ranked voted data and also transformed into a linear assignment problem. Since it becomes a special type assignment problem, we can solve it more efficiently than that of a general linear assignment problem.

As for making our model more realistic and useful, we should introduce a linguistic data as is shown in [Herrera(1996)]. Further in order to validate our model, we should apply the model to assign suitable players to positions using opinions of supervisor and coaches.

References

Borda, J.C.: Memoire sur les elections au scrutiny. Historire de l's Academie de Science, Paris (1781)

Condorcet, M.: Essai sur L's Application de L's Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix, Paris (1785)

Cook, W.D.: Distance-based and at hoc consensus models in ordinal preference ranking. European Journal of Operational Research 172(2), 369–385 (2006)

Cook, W.D., Kress, M.: Relationship between 1 metrics on linear ranking space. SIAM Journal on Applied Mathematics 44(1), 209–220 (1984)

Green, R.H., Doyle, J.R., Cook, W.D.: "Preference voting and project ranking using DEA and cross-evaluation. European Journal of Operational Research 90(6), 461–472 (1996)

Herrera, F., Herrera-Viedma, E., Verdegay, J.L.: A model of consensus in group decision making under linguistic assessments. Fuzzy Sets and Systems 78(1), 73–87 (1996)

Munkres, J.: Algorithms for the assignment and transportation problems. SIAM J. 5(1), 32–38 (1957)

Pentico, D.W.: Assignment problems: A golden anniversary survey. European Journal of Operational Research 176(2), 774–793 (2007)

Tanabe, K., Ishii, H.: "Mathematical approach for consensus formation,". Asia Pacific Management Review 12(3), 145–149 (2007)

Bilevel Toll Optimization Problems: A Heuristic Algorithm Based Upon Sensitivity Analysis

Nataliya I. Kalashnykova, Vyacheslav V. Kalashnikov,
and Roberto C. Herrera Maldonado

Abstract. The authors deal with one of the well-known problems of how to assign an appropriate toll to each toll-arc of a transportation network, which is a combination of both toll and toll-free roads. This task can be formulated as a bilevel programming problem. The top level of such a model is governed by a company that manages the roads (arc tolls) and seeks to increase its profits. At the lower level, there is a group of network users, who make up the demand and look for the routes that minimize their travel costs. In other words, what is sought is a set of tolls that generate the highest revenue for the upper level company, and at the same time, turn out to be attractive for the users. To solve this pricing problem, a direct algorithm based on sensitivity analysis is proposed. In order to make it easier to skip (if necessary) from different pricing environment, that is, from within the vicinity of a local solution to the neighborhood of another, a procedure is proposed making use of the "filled" function method.

1 Introduction

A historical perspective shows the wayforward. In the previous time epochs, the production facilities were established near the centers of consumption because the

Nataliya I. Kalashnykova
Graduate School of Mathematical Sciences, FCFM, UANL,
San Nicolás de los Garza, Nuevo León, Mexico
e-mail: nkalash2009@gmail.com

Vyacheslav V. Kalashnikov · Roberto C. Herrera Maldonado
ITESM, Campus Monterrey, Monterrey, Mexico
e-mail: kalash@itesm.mx

Vyacheslav V. Kalashnikov
Central Economics and Mathematics Institute (CEMI),
Russian Academy of Sciences (RAS), Moscow, Russia

Nataliya I. Kalashnykova · Vyacheslav V. Kalashnikov
Sumy State University, Sumy, Ukraine

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 135–143.
springerlink.com

transportation was expensive, time-consuming and risky. Newly developed transportation systems allowed the producer to enter distant markets, so as to promote economies of scale by increasing sales volume.

An additional consideration is due to the complexity of products, globalization trends, rapid supply chain growth, and that logistics costs have risen. As it was mentioned in [1], according to the data from the IMF (International Monetary Fund), logistics costs represent 12% of gross national product, and at the enterprise level this ranges from 4% to 30% of logistics sales.

Because of this growth, the majority of countries have attached great importance to the development and modernization of their infrastructure to achieve greater participation in the global economy. As we indicated in [1], the Mexican governmental body SCT (Secretaría de Comunicaciones y Transportes), which takes response for the development of infrastructure of communication and transportation, is carrying out the above-mentioned modernization.

Again, according to [1], the land transportation in Mexico, particularly on the highways, accounts for moving 53% of domestic cargos (around 479 million tons) and 98% of passengers (about 3,170 million passengers) who travel across the country, which is why it is considered the backbone of the transportation system.

Moreover, within Mexico, there exists an important road network, classified feeder network, rural and federal. The latter (by the SCT) consists of 19,245 km, of which nearly 40% is made up of concession (toll highways) given to private companies, state governments, or financial institutions. By its total length, this network is one of the largest in the world, only slightly behind France and Italy.

In this paper, we continue considering the problem of assigning optimal tolls to the arcs of a multi-commodity transportation network. The problem was formulated [1] as a bilevel mathematical program where the upper level is managed by a firm that raises revenues from tolls set on (some) arcs of the network, and the lower level was represented by a group of car users traveling along the cheapest paths with respect to a generalized travel cost. The problem can be interpreted as finding equilibrium among tolls generating high revenues and tolls attracting customers.

In [1], the above-mentioned bilevel programming model and the underlying assumptions were widely discussed. Next, in our previous paper [1] an algorithm was proposed and evaluated, which was based on the decreasing of the allowed ranges given by the sensitivity analysis after solving the lower level of the Toll Optimization Problem (TOP). In order to solve this problem efficiently, the authors first reformulated the upper level objective function as the sum of the objective functions for each commodity and then were taking into account the capacity constraints for the arcs.

Bilevel programming offers a convenient framework for modeling the pricing problems, as it allows one to take into account the user's behavior explicitly. One of the simplest instances was analyzed in [2], where a TOP defined over a congestion-free, multi-commodity transportation network was considered. In this setting, a highway authority (the "leader") sets tolls on a subset of arcs of the network, while the users (the "follower") assign themselves to the shortest paths linking their respective origin and destination nodes.

The goal of the leader in this context is to maximize toll revenue. Hence, it is not in its interest to set tolls at too high levels, in which case the users would be discouraged from using the tolled sub-network. The problem can be formulated as a combinatorial program that subsumes NP-hard problems, such as the Traveling Salesman Problem (*see* [3] for a reduction method). Following the initial NP-hardness proof in [3], computational complexity and approximation results have also been obtained in [4].

The aim of the present work is to propose an algorithm based on the allowed ranges resulting from the sensitivity analysis after solving the lower level problem in order to obtain the optimal toll for the highways. This work is a continuation of previous research undertaken by the authors in [1] and [5].

The paper is organized as follows. Section 2 proposes a new bilevel formulation for the Toll Optimization Problem considering arc capacities in the network. Section 3 reduces TOP to a single-level optimization problem by applying certain variational inequalities techniques. Section 4 details the proposed algorithm and provides results of numerical experiments. Finally, Section 5 concludes with plans of further research.

2 The Bilevel Formulation

In this section, we recall the bilevel formulation of the Toll Optimization Problem (TOP) presented in [1] and [5] that takes into account bounded capacities of the arcs of the transportation network. Following mainly [1], TOP can be analyzed as a leader-follower game that takes place on a multi-commodity network $G=(K,N,A)$ defined by a set of origin-destination couples K, a set of nodes N, and a set of arcs A. The latter is partitioned into the subset A_1 of toll arcs and the complementary subset A_2 of toll-free arcs; here $|A_1|=M_1$ and $|A_2|=M_2$, thus yielding $M=M_1+M_2$. We endow each arc $a\in A$ with a fixed travel delay c_a. Also, there is an upper limit capacity q_a associated with each arc $a\in A$ in the network. Toll arcs $a\in A_1$ also involve a toll component t_a to be determined, which, for the sake of consistency, is also expressed in time units. The toll vector $t=\{t_a : a\in A_1\}$ is restricted by the vector $t^{max}=\{t_a^{max} : a\in A_1\}$.

The demand side is represented by numbers n^k denoting the demand for transportation of a commodity $k\in K$ between the origin node $o(k)$ and the destination node $d(k)$. Let us denote the number of commodities as $|K|=r$. A demand vector b^k is associated with each commodity. Its components are defined for every node i of the network as follows:

$$b_i^k = \begin{cases} -n^k, \text{ if } \quad i = o(k), \\ n^k, \quad \text{if } \quad i = d(k), \\ 0, \quad \text{otherwise.} \end{cases} \tag{1}$$

Letting $\left\{x_a^k\right\}_{a\in A}$ denote the set of commodity flows, and i^+ (respectively, i^-) the set of arcs having i as their head node (respectively, tail node), TOP can be formulated as a bilevel programming program (2) – (4):

$$F(x,t) = \sum_{k\in K}\sum_{a\in A_1} t_a x_a^k \to \max_{t,x} \tag{2}$$

subject to

$$0 \le t_a \le t_a^{max} \qquad \forall a \in A_1 \ , \tag{3}$$

and

$$\begin{cases} \varphi_k\left(x^k\right) = \left[\sum_{a\in A_1}\left(c_a + t_a\right)x_a^k + \sum_{a\in A_2} c_a x_a^k\right] \to \min_{x^k}, \\ \text{subject to} \\ \qquad -\sum_{a\in i^-} x_a^k + \sum_{a\in i^+} x_a^k = b_i^k, \qquad \forall i \in N, k \in K, \\ \qquad x_a^k \ge 0, \qquad \forall a \in A, k \in K, \\ \qquad \sum_{k\in K} x_a^k \le q_a, \qquad \forall a \in A. \end{cases} \tag{4}$$

In the above formulation, the leader controls both the toll and flow variables. However the lower 'Argmin' constraint reflects the followers' intention to minimize their total "transportation costs" (in terms of "time delay units" multiplied by the corresponding flow values) under the current toll levels, and subject to the supply requirements. The authors would like to point out that TOP considered in [6] does not include the capacity constraints.

In order to prevent the occurrence of trivial situations, the following conditions are assumed to hold throughout the article (*see* [1], [5] and [6]):

1. There is no profitable toll vector that induces a negative cost (delay) cycle in the network. This condition is clearly satisfied if, for example, all delays c_a are nonnegative.

2. For each commodity, there exists at least one path composed solely of toll-free arcs.

3 The Heuristic Algorithm

In this algorithm, we are going to combine the main structure of the method described in [1] and a new idea consisting in the following. A direct procedure may be represented as the determination of the fastest increase direction for the upper level objective function in terms of the toll variables variations. It is easy to see that the formal "gradient" of this objective function F from (2) is determined by the current total flows along the toll arcs:

$$\frac{\partial F}{\partial t_a}(x,t) = \sum_{k\in K} x_a^k, \quad \forall a \in A_1. \tag{5}$$

We say "formal gradient" because the followers' optimal response is not taken into account in (5). However, as a fastest infinitesimal improvement direction, this vector can be exploited in our heuristic method. The possibility to solve a linear programming problem at the lower level instead of the Nash equilibrium problem (3) – (4) has been justified in the authors' previous papers [1] and [5].

3.1 The Algorithm's Verbal Description

In what follows, we present a rather informal description (a verbal code) of the heuristic method proposed in this paper.

Step 1. Set $i = 0$. Select $t_a^{(i)} = t_a^{min} = 0$ and minimize the lower level (linear) objective function $h_{sum}(x) = \sum_{k\in K}\left(\sum_{a\in A_1}\left(c_a + t_a^{min}\right)x_a^k + \sum_{a\in A_2} c_a x_a^k\right)$ subject to the flow conservation and capacity constraints in order to obtain the optimal response $x\left(t^{(i)}\right)$. Compute the leader's objective function's value $\psi\left(t^{(i)}, x\left(t^{(i)}\right)\right) = \sum_{k\in K}\sum_{a\in A_1} t_a^{(i)} x_a^k\left(t_a^{(i)}\right) = 0$. Go to Step 2.

Step 2. Considering the allowed ranges to stay optimal (ARSO) given by the sensitivity analysis table obtained upon having solved the problem presented in the Initialization, select the (toll-arc) variables $x_a\left(t^{(i)}\right), a \in A_1$, with positive values (that, is all they are basic variables of the current solution of the lower level LP problem). Denote $A_1^+ = \left\{a \in A_1 \,\middle|\, x_a\left(t^{(i)}\right) > 0\right\}$. According to (5), these positive values are (positive) components of the "gradient" vector of the leader's objective function. If $A_1^+ = \varnothing$, then go to Step 4; otherwise, go to Step 3.

Step 3. Compare the positive values of $x_a\left(t^{(i)}\right)$ with those of the maximal allowed increase $\Delta_a^+, a \in A_1^+;$ for each such variable, set

$$t_a^{(i+1)} := \begin{cases} min\left\{t_a^{max}, t_a^{(i)} + min\left\{x_a\left(t^{(i)}\right), \Delta_a^+\right\}\right\}, & \text{if } a \in A_1^+; \\ t_a^{(i)}, \text{ otherwise, i.e.,} & \text{if } a \in A_1 - A_1^+. \end{cases} \tag{6}$$

If $t_a^{(i+1)} > t_a^{(i)}$ for at least one $a \in A_1^+$, then update $i := i+1$ and close the loop by returning to Step 1. Otherwise, i.e. all tolls are the same, go to Step 4.

Step 4. The current set of tolls $\left\{t_a^{(i)}\right\}_{a \in A_1}$ apparently provides for a local maximum of the leader's objective function. In order to try to "skip" to some other possible local maximum solution, apply the "filled" function technique described briefly in subsection 3.2.

Step 5. If, after a number of step 4 repeated, one cannot improve the leader's objective function's value, stop the algorithm, report the current vectors $\left\{t_a^{(i)}\right\}_{a \in A_1}$ and $x\left(t^{(i)}\right)$ as a (global) optimum solution.

3.2 Application of "Filled" Functions

Our heuristic algorithm, based upon sensitivity analysis, also involves application of "filled" function concept first proposed in [7]. This method works, according to studies in [7], under the assumption that a local minimum of a function, which is continuous and differentiable in R^n, has been found. So the aim is to find another (better than the current) local minimum or determine that this is the global minimum of the function within the domain of the closed (polyhedral) set $\Omega \subseteq R^n$. According to [7, 8], a "filled" function can be defined as follows:

Definition 3.1. Let x^* be a local minimum point of a function $f : \Omega \to R$, and $\overline{x}_0 \in \Omega$ be such that $\overline{x}_0 \neq x^*$ and $f\left(\overline{x}_0\right) \leq 5/4 f\left(x^*\right)$. A continuously differentiable function $P_{x^*} : \Omega \to R$ is called a *filled function* for the minimization problem $\min\limits_{x \in \Omega} f(x)$ at x^* with $f\left(x^*\right) > 0$, if

- x^* is a strict local maximizer of $P_{x^*}(x)$ on Ω;
- Any local minimizer $\overline{x}$ of $P_{x^*}(x)$ on Ω starting from $\overline{x}_0$ satisfies $f(\overline{x}) < f\left(x^*\right)/2$ or $\overline{x}$ is a vertex of Ω;

- Any $\tilde{x} \in \Omega$ with $\nabla P_{x^*}(\tilde{x}) = 0$ satisfies $f(\tilde{x}) < f(x^*)/2$;
- Any local minimizer $\hat{x} \in \Omega$ of $f(x)$ on Ω with $f(\hat{x}) \le f(x^*)/4$ is also a local minimizer of $P_{x^*}(x)$ on Ω.

Now what can be undertaken is a description of the steps of the combined direct-and-"filled"-function approach to solve the bilevel TOP.

1. Begin by assigning initial values of zero toll costs, allowing one to relax the problem of the follower and work with a linear programming model at the lower level.
2. After solving the (linear programming) problem of the follower to determine the flow in the arcs and obtaining a value for the leader's objective function, then perform sensitivity analysis for the model of the follower, taking into account only toll-arc variables of the current solution.
3. Now, having listed the possible increases in the coefficients of the objective function of the follower derived from the sensitivity analysis data, and based upon the formal "gradient" vector of the upper level objective function *F* defined by (2), we update the active toll vector $\left\{t_a^{(i)}\right\}_{a \in A_1}$ by formula (6).
4. When positive increments of *t* cannot be obtained anymore based on the sensitivity analysis and the formal "gradient" of function *F*, apply the "filled" function method (*cf.* [7, 8]). A new function is created based on the leader's objective function (that is, the filled function $P_{x^*}(x)$ from Definition 3.1) and a new toll vector is calculated.
5. Once there is a new toll vector, go to Step 2 and close the loop. In case the "filled" function method, after several attempts in row, does not provide a better value for the leader's objective function, this can mean that a global optimum has been reached and the algorithm can stop. The multi-commodity flow corresponding to the final toll values gives the optimal solution for the followers, too.

4 Numerical Results

The authors applied the proposed heuristic algorithm to many numerical examples. However, the paper's volume restrictions allow us to provide only one simple example's description and numerical results.

4.1 The Simple Example

When applied to a simple example (see the graph of the network of highways in Fig. 1 in Appendix A, where only one commodity must be transported in the amount of 10 units from the origin node *O* to the terminal node *T,* and the arc delays are:

$c_{OA} = 15, c_{OB} = 8, c_{OB'} = 13, c_{OE} = 3, c_{ED} = 6, c_{EC} = 9, c_{BC} = 14, c_{AC} = 7, c_{CT} = 10,$
$c_{DT} = 7, c_{AD} = 8, c_{BD} = 11$)

the vector $t = (t_{OE}, t_{DT}, t_{ED}, t_{OB}, t_{AC}, t_{AD})$ starts with zero values. Now, at the lower level, we have a linear programming problem, which we can solve with a software package; we use LINGO12 and MATLAB_R2010a, generating a result for the problem of the followers and Sensitivity Analysis results. Only two iterations were enough to find the optimal solution shown in Table 1.

Table 1 Optimal Solution for the Simple Example

Num. iter. i	*Optimal Tolls*	*Nonzero flows*	*Max. Profit*	*Min. Cost*
2	$t^* = (15,6,0,5,6,1)$	$x^*_{OE} = x^*_{ED} = = x^*_{DT} = 10$	$F^*(x^*, t^*) = 210$	$\varphi(x^*) = 370$

It is interesting to note that such problems usually have a very large number of optimal solutions. For instance, in this simple example, there are 1815 different optimal solutions providing the same maximum profit 210 for the leader.

Numerical experiments were conducted with 16 more examples, their descriptions and the results will be published elsewhere.

5 Conclusion

A heuristic procedure to solve the bilevel Toll Optimization Problem is proposed and justified. The procedure consists in combining the formal "gradient" direction of the upper level objective function and the information supplied by Sensitivity Analysis procedure, in order to determine the toll values increments. The algorithm's work is illustrated by a numerical experiment with a small-sized TOP problem. The future research will be concentrated on revealing a stable procedure of "skipping" to other clusters of paths representing locally optimal solutions, in order to converge to a global solution using the "filled" function method.

Acknowledgements. This work was financially supported by the R&D Department (Cátedra de Investigación en Métodos de Solución de Problemas de Programación Binivel) CAT-174 of the Tecnológico de Monterrey (ITESM), Campus Monterrey, and by the SEP-CONACYT grant CB-2008-01-106664, Mexico. Also, the work of the second author was supported by the SEP-CONACYT grant CB-2010-01-127691.

References

1. Kalashnikov, V.V., Kalashnykova, N.I., Herrera Maldonado, R.C.: Solving Bilevel Toll Optimization Problems by a Direct Algorithm Using Sensitivity Analysis. In: Clute, R.C. (ed.) Proc. of the 2011 New Orleans International Academic Conference, New Orleans, USA, March 21-23, pp. 1009–1018 (2011)

2. Labbé, M., Marcotte, P., Savard, S.: A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing. Management Science 44, 1608–1622 (1998)
3. Marcotte, P., Savard, S., Semet, F.: A Bilevel Programming Approach to the Traveling Salesman Problem. Operations Research Letters 32, 240–248 (2004)
4. Roch, S., Savard, S., Marcotte, P.: Design and Analysis of an Algorithm for Stackelberg Network Pricing. Networks 46, 57–67 (2005)
5. Kalashnikov, V.V., Camacho, F., Askin, R., Kalashnykova, N.I.: Comparison of Algorithms Solving a Bilevel Toll Setting Problem. International Journal of Innovating Computing, Information and Control 6(8), 3529–3549 (2010)
6. Didi-Biha, M., Marcotte, P., Savard, S.: Path-Based Formulationof a Bi-Level Toll Setting Problem. In: Dempe, S., Kalashnikov, V.V. (eds.) Optimization with Multi-Valued Mappings: Theory, Applications and Algorithms, pp. 29–50. Springer Science, Boston (2006)
7. Renpu, G.E.: A Filled Function Method for Finding a Global Minimizer of a Function of Several Variables. Mathematical Programming 46, 191–204 (1988)
8. Wu, Z.Y., Bai, F.S., Mammadov, M., Yang, Y.J.: An Auxiliary Function Method for Systems of Nonlinear Equations. In: Proc. of the 7th International Conference on Optimization: Techniques and Applications (ICOTA 2007). Kobe International Conference Center, Japan, pp. 259–264 (December 2007)

Appendix – Graph of the Simple Example

In this appendix, one can see the graph of the simple example. In the graph, the solid lines correspond to the toll-free arcs and the dotted lines represent the toll arcs. In Fig. 1, only one commodity must be transported in the amount of 10 units from the origin node *O* to the terminal node *T*.

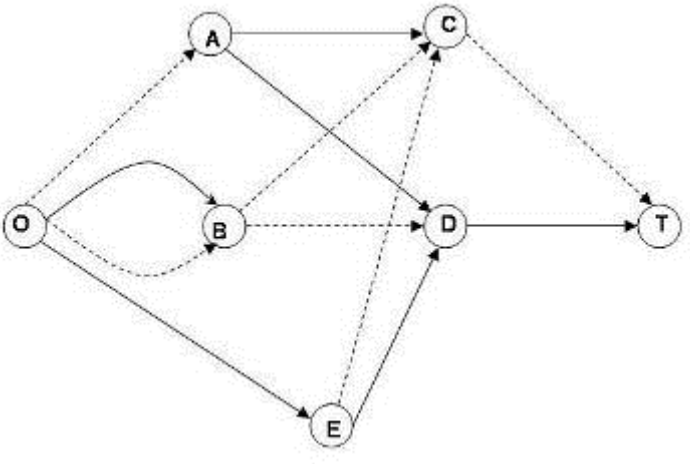


Fig. 1 The simple example

Building a Type II Fuzzy Qualitative Regression Model

Yicheng Wei and Junzo Watada

Abstract. The qualitative regression analysis models quantitatively change in the qualitative object variables by using qualitative values of multivariate data (membership degree or type I fuzzy set), which are given by subjective recognitions and judgments. From fuzzy set-theoretical points of view, uncertainty also exists when associated with the membership function of a type I fuzzy set. It will have much impact on the fuzziness of the qualitative objective external criterion. This paper is trying to model the qualitative change of external criterion's fuzziness by applying type II fuzzy set (we will use type II fuzzy set as well as type II fuzzy data in this paper). Here, qualitative values are assumed to be fuzzy degree of membership in qualitative categories and qualitative change in the objective external criterion is given as the fuzziness of the output.

Keywords: Type II fuzzy qualitative regression model, quantification, type I fuzzy set, type II fuzzy set, type I fuzzy number, type II fuzzy number, linear programming, LP.

1 Introduction

In the real world, people do not often use numerical expressions to judge things. Actually, the recognition, judgment and evaluation activities that humans carry out are commonly expressed in qualitative linguistic terms. For example,when describing features of financial markets, people will use words such as "excellent," "bullish," or "bearish ." However, it won't be easy to compare qualitative judgments without learning the evaluative structure underneath them. If we could replace qualitative expressions with numerical expressions, the problem will be solved easier and better.

Yicheng Wei · Junzo Watada
Graduate School of Information, Production and Systems, Waseda University
2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan
e-mail: weiyicheng0000@gmail.com, junzow@osb.att.ne.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 145–154.
springerlink.com

In order to accomplish this, Chikio Hayashi proposed the quantification theories [1], which include four quantification models [2]: qualitative regression analysis, qualitative discriminate analysis, dual scaling, multi-dimensional scaling. The first quantification method, which is the qualitative regression model, is to find the relationship between qualitative variables and numerical object variable given in the samples. Using values "1" or "0," it quantifies qualitative evaluation by absolute judgment like "Yes" or "No." Computer analysis is therefore easy.

On the other hand, fuzzy set (type I fuzzy set) was introduced in 1965 by Lofti A Zadeh in his paper "Fuzzy Sets" [9]. After that, Tanaka developed fuzzy regression model [4], and Watada and Tanaka expanded fuzzy quantification method in 1987 [8] regarding that the qualitative evaluation in real world will not be absolutely described as yes or no. The handling of type I fuzzy data and fuzzy events provides the basis of fuzzy quantification model. The qualitative input thus will be able to be expressed by any degree within [0,1]. On the background that the membership function of a type-1 fuzzy set has no uncertainty associated with it, Lotfi A. Zadeh invented Type-2 fuzzy sets in 1989 [10]. A type-2 fuzzy set lets us incorporate fuzziness about the membership function into fuzzy set theory, which may be described as "the fuzziness of fuzziness" and is a way to address the above concern of type-1 fuzzy sets head-on.

The qualitative regression analysis models quantitative change in the objective criterion with qualitative values of multivariate data (membership degree) which are given by subjective recognition and judgment. While it does not discuss the uncertainty on the membership function of a type-1 fuzzy set and its impact on membership grade of the objective criterion. Motivated by that, this paper is trying to model the quantitative change of objective criterion's fuzziness by applying type II fuzzy coefficient. Here, qualitative values are assumed to be fuzzy degree of membership in qualitative categories and qualitative change in the objective criterion is given as range of membership degree.

The Hayashi's approach employed in his qualitative regression analysis deals with precise, numeric, statistical data that assumes system coefficients are precise and applied minimum squared method to solve the problem. Although statistical regression has many applications, problems can occur when there is vagueness in the relationship between independent variable and dependent variables, or if there is ambiguity associated with the event. These are the very situations fuzzy regression can address. The first fuzzy linear regression model, which treats fuzzy data instead of statistical data, was proposed by Tanaka et al. [3]. Tanaka and Watada [5] presented possibilistic regression based on the concept of possibility measure.

The objective of this paper is to introduce type II fuzzy numbers to the qualitative regression analysis. The study will use fuzzy possibilistic regression as the modeling structure. The remainder of this paper is organized as follows. In Section 2, we cover some preliminaries of type II fuzzy numbers. Section 3 formulates the Type II Fuzzy Qualitative regression model. And we will give a solution to the problem in Section 4. Finally, concluding remarks are given in Section 5.

2 Fuzzy Numbers

A crisp set is normally defined as follows:

Definition 1. A crisp set is a collection of elements or objects $x \in X$, which can be finite and countable, where X denotes a universal set that consists of all the considered elements. Each single element can either belong to a set A denoted by $A \subseteq X$ or not belong to A written as $A \nsubseteq X$.

In the former case, the statement "x belongs to A" is true, whereas in the latter case this statement is false.

2.1 Fuzzy Sets

Such a classical set can be described in different ways; either one can enumerate the elements that belong to the set, one can describe the set analytically, for instance, one can define the member elements by using the characteristic function.

Definition 2. Characteristic function is a function defined on a set X that indicates membership of an element in a subset A of X, having the value 1 for all elements of A and the value 0 for all elements of X not in A.

$$\mu_A : \ X \to \{0,1\}$$
$$\mu_A = \begin{cases} 1 & \text{if } X \in A \\ 0 & \text{if } X \text{ not } \in A \end{cases}$$

Expanded from the crisp set, for a fuzzy set, the characteristic function allows various degrees of membership for the elements of a given set.

Definition 3. If X is a collection of objects denoted generically by x then a fuzzy set $\tilde{A}$ in X is a set of ordered pairs:

$$A = \{(x, \mu_A(x)) | x \in X\}$$

where μ_A is called the membership function or grade of membership (also degree of compatibility or degree of truth) of x in which maps X to the membership space M.

When M contains only the two points 0 and 1, A is non-fuzzy and μ_A is identical to the characteristic function of a non-fuzzy set. The range of the member-ship function is a subset of the nonnegative real numbers whose supremum is finite. Elements with a zero degree of membership are normally not listed.

Definition 4. The support of a fuzzy set A is the crisp set of all $x \in X$ such that $\mu_A(x) > 0$.

Definition 5. The(crisp)set of elements which belong to the fuzzy set A at least to the degree $\geq \alpha$ is called $\alpha-$ level set and X may be a set of integers I or of real values R.

2.2 Type I Fuzzy Numbers

A type I fuzzy number is a special case of fuzzy set, which is an extension of a regular data in the sense that it does not refer to one single value but rather to a connected set of possible values, where each possible value has its own weight

$$A = \{(x, \mu_A(x)) | x \in X\} \tag{1}$$

which is between 0 and 1.

2.3 Type II Fuzzy Numbers

Expanded from type I fuzzy set, type II fuzzy set expresses the non-deterministic membership degree with imprecision and uncertainty for an element that belongs to a set.

Definition 6. A type II fuzzy set denoted by $\{x, \mu_{\tilde{A}}(x)\}$, is characterized by a type II membership function $\mu_{\tilde{A}}(x, \mu_{\tilde{A}}(x))$,where $x \in X$, $\mu_{\tilde{A}}(x) \subseteq [0,1]$, which is the membership degree of x and $\mu_{\tilde{A}}(x, \mu_{\tilde{A}}(x))$ defined in Equation (3).

$$\tilde{A} = \{x, \mu_{\tilde{A}}(x) | x \in X\} \tag{2}$$

$$\tilde{\tilde{A}} = \{x, \mu_{\tilde{A}}(x), \mu_{\tilde{A}}(x, \mu_{\tilde{A}}(x)) | x \in X\} \tag{3}$$

Definition 7. A type II fuzzy number is a special case of type II fuzzy set, which has convex type II membership function.

Definition 8. $U[\underline{\mu}, \overline{\mu}]$ constructs the support of a fuzzy membership degree for all $\mu \in U$ satisfies that $\mu_{\tilde{A}} > 0$.

3 Fuzzy Qualitative Regression Model

3.1 Type I Fuzzy Qualitative Regression Model

Using fuzzy linear function, we denote the Type I fuzzy regression model as

$$Y_i = A_1 x_{i1} + \cdots + A_j x_{ij} + \cdots + A_K x_{iK} \tag{4}$$

where i denotes sample i for $i = 1, \cdots, n$. Y_i is a grade of a fuzzy sample, which is included in a set of possible real values with a grade Y_i as well as their corresponding membership degree. A_j for $j = 1, 2, \cdots, K$ is a coefficient to describe the membership function of the system for each qualitative attribute.

Sample i has an K-dimensional qualitative attribute (membership degree) input vector $[\mu_A(1),\cdots,\mu_A(K)]$, where

$$\begin{aligned}\mathbf{x}_i = &[(x_{i1},\mu_{x_{i1}}(x_{i1})),\cdots,(x_{ij},\mu_{x_{ij}}(x_{ij})),\cdots,(x_{iK},\mu_{x_{iK}}(x_{iK}))]\\ &\text{for } x_{ij}\in X_i \text{ for } i=1,\cdots,K\\ &0\le\mu_x(x_{ij})\le 1\end{aligned} \tag{5}$$

According to Tanaka's definition, coefficient A_j for sample i, therefore, can be defined as:

$$A_j=[\underline{a_j},\overline{a_j}]=\{A_j:\underline{a_j}\le a_j\le\overline{a_j}\},\qquad j=1,\cdots,n \tag{6}$$

where $\underline{a_j}$ and $\overline{a_j}$ define the support of A_j.

Following Zadeh's extension principle, which enables us to extend any map h: $S_1\times\cdots\times S_K\to S$ where $S_1\times\cdots\times S_K$ and S are sets to map $h^*:F(S_1)\times\cdots\times F(S_K)\to F(S)$, for every set X.

$F(X)$ denotes the class of fuzzy subsets of X; namely the formula to define h^* is

$$h^*(s_1,\cdots,s_n)(x)=\sup\{s_1(x_1)\wedge\cdots,\wedge s_n(x_K)/h(x_1,\cdots,x_K)\}$$

For $s_1\in F(S_1),\cdots,s_n\in F(S_n)$ and $x\in S$, we are able to rewrite the formula as follows:

$$\begin{aligned}\tilde{Y}_i&=\tilde{A}_0+\sum_{j=1}^{K}\tilde{A}_j\mu_i(j)\\ \{\underline{Y_i},\overline{Y_i}\}&=[\underline{a_0},\overline{a_0}]+\sum_{j=1}^{K}[\underline{a_j},\overline{a_j}]\times\mu_{x_{ij}},\end{aligned} \tag{7}$$

where $A^h=[\underline{A^h},\overline{A^h}]$. We also denote $A^h=[\underline{\mu_A^h},\overline{\mu_A^h}]$ for the simplicity.

3.2 The Solution of the Problem

The basic idea of the Tanaka's approach, often referred to as possibilistic regression model, was to minimize the fuzziness of the model by minimizing the total spread of the fuzzy coefficients, subject to including all the given data.

$$\min\ C(=\min\sum_{i=0}^{K}\{\overline{a_i}-\underline{a_i}\}) \tag{8}$$

Note that the observed fuzzy numbers, adjusted for the h-level attribute, which will be discussed in the next section should be contained within the estimated fuzzy output, adjusted for the h-level attribute.

According to Equations (5), (7), (8) and (9), the constraints shall be that theory, we are able to rewrite the formula as follows:

$$
\begin{aligned}
&\underline{a_0}^h + \sum_{j=1}^{K} [\mu_{\tilde{A}}(j) \cdot \underline{a_j}]^h \leq \underline{y_i^h}, \\
&\overline{a_0}^h + \sum_{j=1}^{K} [\mu_{\tilde{A}}(j) \cdot \overline{a_j}]^h \geq \overline{y_i^h}, \\
&\underline{a_j} \leq \overline{a_j}, \\
&\overline{0} \leq \mu_{\underline{\tilde{A}}}(j) \leq 1 \\
&i = 1, \cdots, n; j = 1, \cdots, K, \\
&0 \leq \underline{y_i} \leq \overline{y_i} \leq 1.
\end{aligned}
$$

3.3 The Formulation of Type II Fuzzy Qualitative Regression Model

We have just introduced type I fuzzy qualitative regression model. This section aims to build a new model by applying type II fuzzy numbers. The type I qualitative fuzzy regression has remarked that how much the observed value is belonging to the supposed value by solving the coefficient on system's membership function. And the basic idea behind type II fuzzy qualitative regression model , referred as FQRM-II, is to determine the impact of uncertainty on qualitative membership degree on the objective criterion's fuzziness.

As we have mentioned in 3.1, the observed fuzzy number in type I fuzzy qualitative regression will be the set of real values, which is expressed as $[\underline{y_i}, \overline{y_i}]$. While for FQRM-II, the observed fuzzy data shall be the membership degree of the external criteria, whose value is between 0 and 1 that each observed fuzzy data are denoted in the following:

$$
\mu_{\tilde{Y}_i} = [\underline{\mu_{\tilde{Y}}(i)}, \overline{\mu_{\tilde{Y}}(i)}], \qquad \mu_{\tilde{Y}}(j) \in [0,1] \tag{9}
$$

which we may define as qualitative observed outputs . On the other hand, the definition of type II fuzzy data allows a type II function to describe the uncertainty of system membership degree for each input's membership degree for the observed data. And the coefficients express the fuzziness of the given input's membership degree. Thus, for each $\mu_{\tilde{x}}(j)$, there may exist

$$
\tilde{\tilde{A}} = \{x_j, \mu_{\tilde{A}}(x_j), \mu_{\tilde{\tilde{A}}}(\mu_{\tilde{A}}(x_j))\}, \tag{10}
$$

where

$$
[\underline{\mu_{\tilde{A}}(j)}, \overline{\mu_{\tilde{A}}(j)}]. \tag{11}
$$

The problem then can be formulated as

$$
\mu_{\tilde{Y}_i} = \sum_{j=0}^{K} \{\tilde{\tilde{A}}_j \times \mu_{\tilde{x}_{ij}}\}, \tag{12}
$$

where i is the i-th sample, $i = 1,\cdots,n$; j is the j-th attribute, $j = 1,\cdots,K$ and $\mu_{\tilde{Y}_i} = [\underline{\mu_{\tilde{Y}}(i)}, \overline{\mu_{\tilde{Y}}(i)}]$, while $\mu_{\tilde{x}_{ij}} = [\underline{\mu_{\tilde{x}}(j)}, \overline{\mu_{\tilde{x}}(j)}]$. Each of the coefficient can be expressed by $\tilde{\tilde{A}} = [\underline{\mu_{\tilde{A}}(j)}, \overline{\mu_{\tilde{A}}(j)}]$. In later passage, we may take $[\underline{y_i}, \overline{y_i}]$, $[\underline{x_{ij}}, \overline{x_{ij}}]$ instead of $[\underline{\mu_{\tilde{x}}(j)}, \overline{\mu_{\tilde{x}}(j)}]$, $[\underline{\mu_{\tilde{Y}}(j)}, \overline{\mu_{\tilde{Y}}(j)}]$ in order to simplify the expression.

3.3.1 h-Level Attribute

Theoretically, the system's supports should be sufficient enough to include all the observed data points of the sample. Thus, the least observed one and the biggest one will construct the support of output membership function. Whereas in practice, we may add some confidence in out-of-sample projection using the estimated fuzzy regression model. And we will realize the concern by extending the supports. We may use Definition 2 to fix a h level and make sure that membership degree of all observed data is above or equal to h.

Assume that $\tilde{Y}$ follows a membership function $\mu_{\tilde{\tilde{Y}}}$. According to the convexity of $\mu_{\tilde{\tilde{Y}}}$, we may have two values for an assigned h

$$h = \tilde{\tilde{Y}}(\underline{\mu_{\tilde{A}}(j)}^h) = \tilde{\tilde{Y}}(\overline{\mu_{\tilde{A}}(j)}^h)$$

We can easily get values of $\underline{\mu_{\tilde{A}}(j)}^h)$ and $\overline{\mu_{\tilde{A}}(j)}^h)$ through calculating $\tilde{\tilde{Y}}^{-1}(h)$ The observed fuzzy data shall fall within the interval $[\underline{\mu_{\tilde{A}}(j)}^h), \overline{\mu_{\tilde{A}}(j)}^h)]$. It is obvious that an h-certain attribute decides the strictness of the model.

We may assign the h-level attribute conception to FQRM-II as well. What is different to the situation we discussed in last section, there will be three situations for range of outcome by multiplying input and coefficient's edge

$$\begin{aligned}
&(1)\quad \overline{a_j}^h \geq \underline{a_j}^h \geq 0 \quad &&: (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\underline{a_j^h} \cdot \underline{x_{ij}^h}, \overline{a_j^h} \cdot \overline{x_{ij}^h}]\\
&(2)\quad \overline{a_j}^h \leq \underline{a_j}^h \leq 0 \quad &&: (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\overline{a_j^h} \cdot \overline{x_{ij}^h}, \underline{a_j^h} \cdot \underline{x_{ij}^h}]\\
&(3)\quad \underline{a_j}^h \leq 0 \leq \overline{a_j}^h \quad &&: (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\underline{a_j^h} \cdot \overline{x_{ij}^h}, \overline{a_j^h} \cdot \underline{x_{ij}^h}]\\
& && \quad i = 1,2,\cdots,n;\ j = 1,2,\cdots,K
\end{aligned}$$

Given that we will not be able to choose which situation to apply at first, thus, we must enumerate the constraints by using traditional vector method in order to solve the problem

$$\begin{aligned}
&\min\ C \qquad (= \min\ \textstyle\sum_j \{\overline{a_j} - \underline{a_j}\})\\
&\text{Subject to}\\
&\qquad \underline{\underline{y_1^h}} \geq \underline{\underline{a_1^h}} \cdot \underline{\underline{x_{11}^h}} + \underline{\underline{a_2^h}} \cdot \underline{\underline{x_{12}^h}} + \cdots + \underline{\underline{a_K^h}} \cdot \underline{\underline{x_{1K}^h}}\\
&\qquad \overline{\overline{y_1^h}} \leq \overline{\overline{a_1^h}} \cdot \overline{\overline{x_{11}^h}} + \overline{\overline{a_2^h}} \cdot \overline{\overline{x_{12}^h}} + \cdots + \overline{\overline{a_K^h}} \cdot \overline{\overline{x_{1K}^h}}\\
&\qquad \vdots\\
&\qquad \underline{\underline{y_n^h}} \geq \underline{\underline{a_1^h}} \cdot \underline{\underline{x_{n1}^h}} + \underline{\underline{a_2^h}} \cdot \underline{\underline{x_{n2}^h}} + \cdots + \underline{\underline{a_K^h}} \cdot \underline{\underline{x_{nK}^h}}\\
&\qquad \overline{\overline{y_n^h}} \leq \overline{\overline{a_1^h}} \cdot \overline{\overline{x_{n1}^h}} + \overline{\overline{a_2^h}} \cdot \overline{\overline{x_{n2}^h}} + \cdots + \overline{\overline{a_K^h}} \cdot \overline{\overline{x_{nK}^h}}\\
&\qquad\qquad \text{for } i = 1,\cdots,n;\ j = 1,\cdots,K
\end{aligned} \tag{13}$$

There are $2K$ inequalities for each sample i. Therefore, we will have $K+2*N*2K$ inequalities in total.Unfortunately this problem cannot be solved within a reasonable computing time when K becomes even moderately large. For example, when we have 100 features and $10,000$ samples, the linear programming problem will come with $2*10,000*2^{100}$ constraints and 100 non-negative constraints. Given this, we have to resort to some heuristic strategies, which is more efficient.

4 A Solution Based on Heuristic Method

Given we are not able to decide whether the coefficient is positive or negative at first. We may introduce a trial and error method to approach the consequence. The basic idea behind it is to eliminate the error of estimation on the polar of coefficient by checking the consistency of it. Once the polar of each coefficient becomes consistent, then we will take the outcome as the consequence. Besides, we may introduce a attribute L to accommodate the accuracy of the consequence Consider

$$X_i = [\underline{x_{ij}}, \overline{x_{ij}}]$$

We may define:

$$x_{ij} = \frac{\underline{x_{ij}} + \overline{x_{ij}}}{2}$$

as the approximately estimation of the membership degree we will take x_{ij} as the first input to the model to roughly determine whether the coefficient is positive or not. In the constraints, we make all values included in the upper and lower boundary.

LP problem is described as follows:

$$\begin{aligned}
&\min\ C \qquad (= \min \sum_{j=0}^{K} \{\overline{a_j} - \underline{a_j}\}) \\
&\text{Subject to} \\
&\qquad \underline{\underline{y_1^h}} \geq \underline{\underline{a_1^h}} \cdot x_{11}^h + \underline{\underline{a_2^h}} \cdot x_{12}^h + \cdots + \underline{\underline{a_K^h}} \cdot x_{1K}^h \\
&\qquad \overline{y_1^h} \leq \overline{a_1^h} \cdot x_{11}^h + \overline{a_2^h} \cdot x_{12}^h + \cdots + \overline{a_K^h} \cdot x_{1K}^h \\
&\qquad \vdots \\
&\qquad \underline{\underline{y_n^h}} \geq \underline{\underline{a_1^h}} \cdot x_{n1}^h + \underline{\underline{a_2^h}} \cdot x_{n2}^h + \cdots + \underline{\underline{a_K^h}} \cdot x_{nK}^h \\
&\qquad \overline{y_n^h} \leq \overline{a_1^h} \cdot x_{n1}^h + \overline{a_2^h} \cdot x_{n2}^h + \cdots + \overline{a_K^h} \cdot x_{nK}^h \\
&\qquad \overline{a_j^h} \geq \underline{a_j^h}, \qquad j = 1,2,\cdots,K \\
&\qquad x_{ij} \geq \overline{0} \qquad i = 1,2,\cdots,n;\ \ j = 1,2,\cdots,K
\end{aligned} \tag{14}$$

Through solving the LP problem, we will get the first round's $\overline{a_j(1)}$ and $\underline{a_j(1)}$. At this time, $L = 1$. Notice that the first round is not avoidable given we need use the outcomes to check next round's consistency of the coefficient.

Consider a jth-coefficient if $\overline{a_j(1)^h} \geq \underline{a_j(1)^h} \geq 0$, we will assign $(\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h$ as

$$(\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\underline{a_j^h} \cdot \underline{x_{ij}^h}, \overline{a_j^h} \cdot \overline{x_{ij}^h}]$$

As well as

$$\begin{array}{lll}
(1) & \overline{a_j}^h \geq \underline{a_j}^h \geq 0 & : (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\underline{a_j^h} \cdot \underline{x_{ij}^h}, \overline{a_j^h} \cdot \overline{x_{ij}^h}] \\
(2) & \overline{a_j}^h \leq \underline{a_j}^h \leq 0 & : (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\overline{\underline{a_j^h}} \cdot \overline{\underline{x_{ij}^h}}, \overline{a_j^h} \cdot \underline{x_{ij}^h}] \\
(3) & \underline{a_j}^h \leq 0 \leq \overline{a_j}^h & : (\tilde{\tilde{A}}_j, \mu_{\tilde{x}_{ij}})^h = [\underline{a_j^h} \cdot \overline{x_{ij}^h}, \overline{a_j^h} \cdot \overline{\underline{x_{ij}^h}}] \\
 & & i = 1, 2, \cdots, n; \;\; j = 1, 2, \cdots, K
\end{array}$$

By doing so we have formulate another LP problem ,which is the second round

$$\begin{array}{ll}
\min\ C & (= \min \sum_{j=0}^{K} \{\overline{a_j} - \underline{a_j}\}) \\
\text{Subject to} & \\
 & \overline{\underline{y_1^h}} \geq \overline{\underline{a_1^h}} \cdot x_{11}^h + \overline{\underline{a_2^h}} \cdot x_{12}^h + \cdots + \overline{\underline{a_K^h}} \cdot x_{1K}^h \\
 & \overline{y_1^h} \leq \overline{a_1^h} \cdot x_{11}^h + \overline{a_2^h} \cdot x_{12}^h + \cdots + \overline{a_K^h} \cdot x_{1K}^h \\
 & \vdots \\
 & \overline{\underline{y_n^h}} \geq \overline{\underline{a_1^h}} \cdot x_{n1}^h + \overline{\underline{a_2^h}} \cdot x_{n2}^h + \cdots + \overline{\underline{a_K^h}} \cdot x_{nK}^h \\
 & \underline{y_n^h} \leq \overline{a_1^h} \cdot x_{n1}^h + \overline{a_2^h} \cdot x_{n2}^h + \cdots + \overline{a_K^h} \cdot x_{nK}^h \\
 & a_j^h \geq \underline{a_j^h}, \qquad j = 1, 2, \cdots, K \\
 & x_{ij} \geq \overline{0} \qquad i = 1, 2, \cdots, n; \;\; j = 1, 2, \cdots, K
\end{array} \tag{15}$$

After solving the LP problem above we may get the value of $\overline{a_j}(2)$ and $\underline{a_j}(2)$. At this time, $L = 2$.

If it satisfies

$$\begin{array}{l}
\overline{a_j}(1) \times \overline{a_j}(2) \geq 0 \\
\underline{a_j}(1) \times \underline{a_j}(2) \geq 0
\end{array} \tag{16}$$

Then, it is possible for us to judge that the polar of coefficients has become consistent. Hence, we may take $\overline{a_j}(2)$ and $\underline{a_j}(2)$ as the final outcome in this case. If the condition is not satisfied, the procedure of iteration will be continued till we get both

$$\begin{array}{l}
\overline{a_j}(L-1) \times \overline{a_j}(L) \geq 0 \\
\underline{a_j}(L-1) \times \underline{a_j}(L) \geq 0
\end{array} \tag{17}$$

Besides, we will assign a L0 (which is set as description of the accuracy) for the problem at very be-ginning. After the polar has become consistent , we are allowed to repeat the pro-cedure of iteration till the outcome meets our requirement of accuracy.

5 Concluding Remarks

In this paper, we built a type II fuzzy qualitative regression model by employing type II fuzzy numbers, which discuss the impact of fuzzy degree of membership in qualitative categories on fuzziness of the objective criterion. The proposed model generalized our previous work. And we will continue working on its empirical application .

Acknowledgements. This work is partially supported by Grants-in-Aid for Scientific Research, MEXT (No. 23500289).

References

1. Hayashi, C.: On the Quantification of Qualitative Data from the Mathematico- Statistical Point of View. Annals of the Institute of Statistical Mathematics 2(1), 35–47 (1950)
2. Hayashi, C.: On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathmatico-Statistical Point of View. Annals of the Institute of Stat. Math. 3, 69–98 (1952)
3. Tanaka, H., Asai, K.: Fuzzy linear programming problems with fuzzy number. Fuzzy Sets and Systems 13(1), 1–10 (1984)
4. Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. IEEE Trans. on Systems, Man, and Cybernetics, SMC-12(6), 903–907 (1982)
5. Tanaka, H., Watada, J.: Possibilistic linear systems and their application. International Jounal of Fuzzy Sets and Systems 27(3), 275–289 (1988)
6. Watada, J., Pedrycz, W.: A Possibilistic Regression Approach to Acquisition of Linguistic Rules. In: Pedrycz, W., Skowron, A., Kreinovich, V. (eds.) Handbook on Granular Commutation, ch. 32, pp. 719–740. John Wiley and Sons, Chichester (2008)
7. Watada, J., Wang, S., Pedrycz, W.: Building Confidence-Interval-Based Fuzzy Random. Regression Models. IEEE Transactions on Fuzzy Systems 17(6), 1273–1283 (2009)
8. Watada, J., Tanaka, H.: Fuzzy quantification methods. In: Proceedings, the 2nd IFSA Congress, pp. 66–69 (1987)
9. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
10. Zadeh, L.A.: The Concept of a Linguistic Variable and Its Application to Approximate Reasoning-1. Information Sciences 8(3), 199–249 (1975)

Building Fuzzy Random Autoregression Model and Its Application

Lu Shao, You-Hsi Tsai, Junzo Watada, and Shuming Wang

Abstract. The purpose of economic analysis is to interpret the history, present and future economic situation based on analyzing economical time series data. The autoregression model is widely used in economic analysis to predict an output of an index based on the previous outputs. However, in real-world economic analysis, given the co-existence of stochastic and fuzzy uncertainty, it is better to employ a fuzzy system approach to the analysis. To address regression problems with such hybridly uncertain data, fuzzy random data are introduced to build the autoregression model. In this paper, a fuzzy random autoregression model is introduced and to solve the problem, we resort to some heuristic solution based on σ-confidence intervals. Finally, a numerical example of Shanghai Composite Index is provided.

Keywords: fuzzy autoregression model, fuzzy random data, fuzzy random autoregression model, time series data, confidence interval.

1 Introduction

Many econometric models have been proposed to evaluate different kinds of economic structures and problems. Classical regression model leads to effective statistical analysis of numeric and precise data. Although an autoregression model can reflect the reality to some extent, it is not suitable for situations where the quantities are not functionally related.

Economic analysis faces a very complex economic system which includes substantial uncertainty affected by many factors, for example human behaviors [1]. Because of the characteristics of economic analysis, the incorporation of the concept

Lu Shao · You-Hsi Tsai · Junzo Watada · Shuming Wang
Graduate School of Information, Production and Systems, Waseda University;
2-7 Hibikino, Wakamatsu, Kitakyushu 808-0135 Japan
e-mail: skyjustsl@yahoo.com.cn, tsaikath@gmail.com
junzow@osb.att.ne.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 155–164.
springerlink.com

of fuzzy sets into the statistical models tends to be an efficient approach. After L.A. Zadeh proposed fuzzy set theory [2] in 1965, it is possible to deal with qualitative data such as linguistic data easily by using the definition of member-ship functions. In 1996, Ozawa et al. proposed a fuzzy autoregression model [3] that expresses the possibilities of fuzzified difference sequences. The concept of fuzzy random variable was introduced by Kwakernaak [4], who defined these variables as a measurable function linking a probability space with a collection of fuzzy numbers. To deal with fuzzy random variables, a series of optimization models have been proposed, which help cope with uncertainty due to both randomness and fuzziness, such as fuzzy random goal programming [5], fuzzy random linear programming [6] and fuzzy random reliability optimization models [7].

In this research, the economic data in a certain period of time can be regarded as fuzzy random variables and applied to the fuzzy autoregression model. This study focuses on the generalization of the fuzzy autoregression model and an expected approach to fuzzy random data. To solve the optimization of fuzzy random autoregression model, a realistic heuristic method based on confidence interval is explained. The confidence interval is defined by the expected value and variance of a fuzzy random variable.

The structure of this paper is organized as follows: In Section 2, the definition of fuzzy random variable is introduced. Section 3 formulates the fuzzy random autoregression model and a numerical example of Shanghai Stock Composite Index is given in Section 4. At last, results and conclusions are made in Section 5.

2 Fuzzy Random Variables

Given some universe Γ, let *Pos* be a possibility measure defined on the power set $P(\Gamma)$ of Γ. Let R be the set of real numbers. A function $Y : \Gamma \to R$ is said to be a fuzzy variable defined on Γ (see [8]). The possibility distribution μ_Y of Y is defined by $\mu_Y(t) = Pos\{Y = t\}$, $t \in R$, which is the possibility of event $\{Y = t\}$. For fuzzy variable Y, with possibility distribution μ_Y, the possibility, necessity, and credibility of event $\{Y \leq r\}$ are given as follows:

$$\begin{aligned} Pos\{Y \leq r\} &= \sup_{t \leq r} \mu_Y(t) \\ Nec\{Y \leq r\} &= 1 - \sup_{t > r} \mu_Y(t) \\ Cr\{Y \leq r\} &= \frac{1}{2}(1 + \sup_{t \leq r} \mu_Y(t) - \sup_{t > r} \mu_Y(t)) \end{aligned} \tag{1}$$

From Equation (1), we note that the credibility measure is an average of the possibility and the necessity measures, i.e., $Cr\{\cdot\} = \dfrac{(Pos\{\cdot\} + Nec\{\cdot\})}{2}$. The motivation behind the introduction of the credibility measure is to develop a certain measure, which is a sound aggregate of the two extreme cases, such as the possibility (that expresses a level of overlap and is highly optimistic in this sense) and necessity (that

articulates a degree of inclusion and is pessimistic in its nature). Based on credibility measure, the expected value of fuzzy variable is presented as follows.

Definition 1. (see [9]): Let Y be a fuzzy variable. The expected value of Y is defined as

$$E[Y] = \int_0^{\infty} Cr\{Y \geq r\}dr - \int_{-\infty}^{0} Cr\{Y \leq r\}dr \tag{2}$$

under the condition that the two integrals are finite.

Assume that $Y = (c, a^l, a^r)_T$ is a triangular fuzzy variable whose possibility distribution is given by

$$\mu_{Y(t)} = \begin{cases} \dfrac{x - a^t}{c - a^l}, & a^l \leq x \leq c \\ \dfrac{a^r - x}{a^r - c}, & c \leq x \leq a^r \\ 0, & otherwise \end{cases} \tag{3}$$

Making use of Equation(2), we determine the expected value of Y to be

$$E[Y] = \frac{a^l + 2c + a^r}{4}. \tag{4}$$

Definition 2. (see[10]): Suppose that (Ω, Σ, Pr) is a probability space and F_V is a collection of fuzzy variables defined on possibility space $(\Gamma, P(\Gamma), Pos)$. A fuzzy random variable is a mapping $X : \Omega \to F_V$ such that for any Borel subset B of R, $Pos\{X(\omega) \in B\}$ is a measurable function of ω.

Let X be a fuzzy random variable on Ω. From the previous definition, we know, for each $\omega \in \Omega$, that $X(\omega)$ is a fuzzy variable. Furthermore, a fuzzy random variable X is said to be positive if, for almost every ω, fuzzy variable $X(\omega)$ is positive almost surely.

For any fuzzy random variable X on Ω, for each $\omega \in \Omega$, the expected value of the fuzzy variable $X(\omega)$ is denoted by $E[X(\omega)]$, which has been proved to be a measurable function of ω (see [10], Th.2), i.e., it is a random variable. Given this, the expected value of the fuzzy random variable X is defined as the mathematical expectation of the random variable $E[X(\omega)]$.

Definition 3. (see[10]): Let X be a fuzzy random variable defined on a probability space (Ω, Σ, Pr). Then, the expected value of X is defined as

$$E[\xi] = \int_{\Omega} [\int_0^{\infty} Cr\{\xi(\omega) \geq r\}dr - \int_{-\infty}^{0} Cr\{\xi(\omega) \leq r\}dr]Pr(\omega) \tag{5}$$

Definition 4. (see [10]): Let X be a fuzzy random variable defined on a probability space (Ω, Σ, Pr) with expected value e. Then, the variance of X is defined as

$$Var[X] = E[(X - e)^2] \tag{6}$$

where $e = E[X]$ is given by Definition 3.

3 Fuzzy Random Autoregression Model

3.1 Fuzzy Autoregression Model

A fuzzy autoregression model is defined with the following equations.

$$\begin{aligned} \widetilde{Z}_t &= \phi_1 Z_{t-1} + \cdots + \phi_k Z_{t-k} + u \\ Z_t &\subseteq \widetilde{Z}_t, \widetilde{Z}_t = (\widetilde{\alpha}_t, \widetilde{\beta}_t, \widetilde{\delta}_t) \end{aligned} \tag{7}$$

It is clear that the following relations hold from Equations (7).

$$\begin{aligned} \alpha_t &\geq \widetilde{\alpha}_t \\ \delta_t &\leq \widetilde{\delta}_t \end{aligned}$$

Namely, a value in terms of estimating the fuzzy time-series model includes all of the fuzzy data of the original series. The autoregression parameters $\phi_1, \cdots, \phi_k$ have real values and show the degree that the fuzzy time-series data depend on the past. An error term is the constant term that is characteristic of the model and shows the part of the fuzzy data that does not depend on past data. This is also defined as a triangular fuzzy number $u = (u_\alpha, u_\beta, u_\delta)$.

A fuzzy autoregression model results in the linear programming that minimizes that ambiguity of the model according to the inclusion condition (7) as follows:

$$\begin{aligned} & min \sum_{t=k+1}^{n} (\widetilde{\delta}_t - \widetilde{\alpha}_t) \\ \text{Subject to} \quad & \\ & \alpha_t \geq \widetilde{\alpha}_t = \phi_1 \alpha_{t-1} + \cdots + \phi_k \alpha_{t-k} + u_\alpha \\ & \delta_t \leq \widetilde{\delta}_t = \phi_1 \delta_{t-1} + \cdots + \phi_k \delta_{t-k} + u_\delta \\ & u_\alpha \leq u_\beta \end{aligned} \tag{8}$$

3.2 Fuzzy Random Autoregression Model

Table 1 illustrates a format of data to be dealt with here, where input data and output data Y_{t-n}, for all $n = 0, \cdots, N$ are fuzzy random variables, which are defined as

$$Y_t = \bigcup_{m=1}^{M} \left((Y_t^m, Y_t^{m,l}, Y_t^{m,r})_\tau, p_t^m \right) \tag{9}$$

This means that all values are given as fuzzy numbers with probabilities.

Let us denote fuzzy linear regression model as follows:

$$\widetilde{Y}_t = \varphi_1 Y_{t-1} + \cdots + \varphi_k Y_{t-k} + u \tag{10}$$

Table 1 Fuzzy random input-output data

Sample	Output	Input			
0	Y_t	Y_{t-1},	Y_{t-2},	$\cdots$,	Y_{t-k}
1	Y_{t-1}	Y_{t-2},	Y_{t-3},	$\cdots$,	$Y_{t-(k+1)}$
2	Y_{t-2}	Y_{t-3},	Y_{t-4},	$\cdots$,	$Y_{t-(k+2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
N	Y_{t-N}	$Y_{t-(N+1)}$,	$Y_{t-(N+2)}$,	$\cdots$,	$Y_{t-(N+k)}$

where $\widetilde{Y}_t$ denotes an estimate of the output and $varphi_1, \cdots, \varphi_k$ are parameters which have real values when triangular fuzzy random data Y_{t-k} as shown in Table 1.

When outputs Y_{t-k} are given at the same time, we can determine the fuzzy random linear model so that the model includes all given fuzzy random outputs. Therefore, the following relation should hold:

$$\widetilde{Y}_t = (\widetilde{\alpha}_t, \widetilde{\beta}_t, \widetilde{\delta}_t) \supset_{FR} Y_t \tag{11}$$

where $\supset_{FR}$ is a fuzzy random inclusion relation whose precise meaning will be explained later on. Following the principles of fuzzy arithmetic, a fuzzy random autoregression model results in the linear programming that minimizes that ambiguity of the model according to the inclusion condition (7) as follows:

$$\begin{aligned} & min \sum (\widetilde{\delta}_t - \widetilde{\alpha}_t) \\ \text{Subject to} & \\ & \alpha_t \geq \widetilde{\alpha}_t \\ & \delta_t \leq \widetilde{\delta}_t \\ & u_\alpha \leq u_\beta \end{aligned} \tag{12}$$

In this study, the confidence-interval based inclusion, which combines the expectation and variance of fuzzy random variables is employed to deal with the model(12). There are also some other ways to define the fuzzy random inclusion relation $\supset_{FR}$, which will yield more complicated fuzzy random regression models.

Before building the fuzzy random regression model, we define the confidence interval which is induced by the expectation and variance of a fuzzy random variable. When we consider the one sigma confidence $(1 \times \sigma)$ interval of each fuzzy random variable, we can express it as the following interval

$$I[e_X, \sigma_x] = [E(X) - \sqrt{Var(X)}, E(X) + \sqrt{Var(X)}] \tag{13}$$

which is called a one-sigma confidence interval. Similarly, we can define two sigma and three-sigma confidence intervals. All of these confidence intervals are called σ-confidence intervals. Then α_t and δ_t can be obtained by:

$$\begin{aligned} \alpha_t &= E(X) - \sqrt{Var(X)} \\ \delta_t &= E(X) + \sqrt{Var(X)} \end{aligned} \tag{14}$$

which will be used in the calculation of the fuzzy random autoregression model(12).

4 Application to Analyzing Stock Price Time Series Data

In this section, we employ the Shanghai Composite Index data which indicates the trend of the whole stock market in China. We use the daily close data from April 11th to June 3rd, 2011. The data of the transaction days in one week are formulated to fuzzy data shown in Table 2 and regarded as fuzzy random data shown in Table 3.

Table 2 Fuzzy data

Time	Fuzzy variables
Week1	[3020, 3035, 3050]
Week2	[3000, 3028.5, 3057]
Week3	[2886, 2925, 2964]
Week4	[2863, 2897.5, 2932]
Week5	[2845, 2867.5, 2890]
Week6	[2848, 2860, 2872]
Week7	[2709, 2742, 2775]
Week8	[2705, 2724.5, 2744]

Table 3 Fuzzy random data

Time	Fuzzy variables	Probability 1	Fuzzy variables	Probability 2
Week 1	[3030, 3020, 3040]	0.4	[3040, 3030, 3050]	0.6
Week 2	[3019, 3000, 3038]	0.8	[3038, 3019, 3057]	0.2
Week 3	[2912, 2886, 2938]	0.5	[2938, 2912, 2964]	0.5
Week 4	[2886, 2863, 2909]	0.75	[2909, 2886, 2932]	0.25
Week 5	[2860, 2845, 2875]	0.2	[2875, 2860, 2890]	0.8
Week 6	[2856, 2848, 2864]	0.8	[2864, 2856, 2872]	0.2
Week 7	[2731, 2709, 2753]	0.5	[2753, 2731, 2775]	0.5
Week 8	[2718, 2705, 2731]	0.4	[2731, 2718, 2744]	0.6

By the definition introduced in Section 2, we can get the expectation and standard deviation as well as the confidence intervals of the data as shown in Table 4 and Table 5.

Table 4 Expectation and standard deviation of fuzzy random data

Time	Expectation	Standard Deviation
Week 1	3036	7.4
Week 2	3022.8	12.4
Week 3	2925	19.5
Week 4	2891.8	15.7
Week 5	2872	9.8
Week 6	2857.6	5.2
Week 7	2742	16.5
Week 8	2725.8	7.9

Table 5 Confidence intervals of fuzzy random data

Time	Confidence intervals
Week1	[3028.6, 3043.4]
Week2	[3010.4, 3035.2]
Week3	[2905.5, 2944.5]
Week4	[2875.3, 2906.7]
Week5	[2862.2, 2881.8]
Week6	[2852.4, 2862.8]
Week7	[2725.5, 2758.5]
Week8	[2717.9, 2733.7]

4.1 Fuzzy Autoregression Model

We employ a calculus of finite differences to filter out the time-series trend data, which enables us to use the first-order difference-equation to write the following:

$$Z_t = (Z_t^L, Z_t^C, Z_t^U = (min(Y_t - Y_{t-1}, Y_t^C - Y_{t-1}^C, max(Y_t - Y_{t-1})) \tag{15}$$

Then the fuzzy random autoregression model is built as follows:

$$\begin{aligned} & min \sum_{t=3}^{7} (\widetilde{\delta}_t - \widetilde{\alpha}_t) \\ \text{Subject to} \quad & \\ & \alpha_t \geq \widetilde{\alpha}_t = \phi_1 \alpha_{t-1} + \phi_2 \alpha_{t-2} + u_\alpha \\ & \delta_t \leq \widetilde{\delta}_t = \phi_1 \delta_{t-1} + \cdots + \phi_2 \delta_{t-2} + u_\delta \\ & u_\alpha \leq u_\beta \end{aligned} \tag{16}$$

We analyze the data by employing the fuzzy autoregression model with the triangular fuzzy number. The coefficients of this model are determined as follows:

$$Z_t = -0.616 Z_{t-1} - 0.472 Z_{t-2} + (-229.9, -82.3, 65.3) \tag{17}$$

An original series and estimated series are shown in Figure 1.

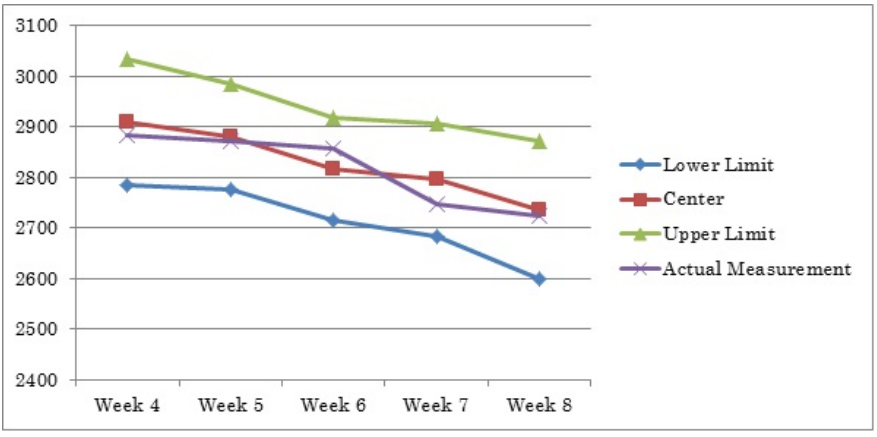


Fig. 1 The conjectured result by the fuzzy autoregression model

Figure 1 shows that the estimated model has a large width of possibility. Numerically, the width of the possibility of the model is 231.1 on average, 272.5 at the maximum and 201 at the minimum. The ambiguity of this model is extremely large.

4.2 Fuzzy Random Autoregression Model

Next, we analyze the Shanghai Composite Index by employing the fuzzy random autoregression model which is proposed in this paper. Similar to the case of the fuzzy autoregression model, the first-order difference series is taken to filter out the time-series trend data.

In the estimated fuzzy random autoregression model, the coefficients are determined as follows:

$$Z_t = -0.763Z_{t-1} - 0.733Z_{t-2} + (-192.34, -92.95, 6.44) \tag{18}$$

The model that is obtained by the fuzzy random autoregression model has negative coefficients which are the same as the result obtained by the fuzzy auto-regressive model.

The original series and estimated series are shown in Figure 2.

As shown in Figure 2, the estimated model has a small width and results in the low level of fuzziness. Numerically, the width of the possibility of the model is 147.46, 176.7 at the maximum and 127.9 at the minimum. In Figure 2, the width of the model is smaller than that in Figure 1. The central value of the estimated value shows a value that is almost the same as the original series.

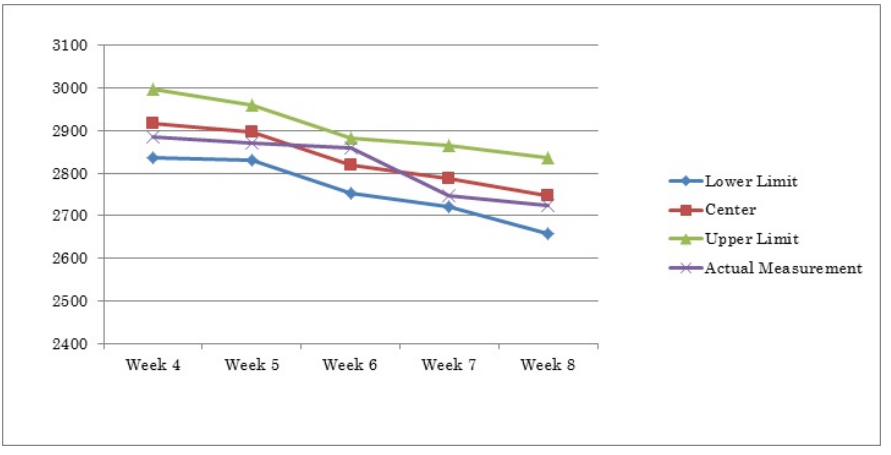


Fig. 2 The conjectured result by the fuzzy random autoregression model

5 Results and Conclusions

Let us compare the results between a fuzzy autoregression model and a fuzzy random autoregression model employing the triangular fuzzy number, which are illustrated in Figure 1 and Figure 2. These figures show that the fuzzy autoregression mode l is estimated as the model with a large width of possibility and the fuzzy random autoregression model is estimated as the model with a small width. The ambiguity of the fuzzy random autoregression model is little, and this model estimated the original series more correctly. However, the fuzzy autoregression model includes the entire fluctuation of an original series by estimating its fluctuation with a large value.

From the results of the fuzzy random autoregression model, we can see that the central value of the estimated series is quite near to the actual series, which means the fuzzy random autoregression model we built can perform well in fore-casting the future data of the time series. The forecasted confidence intervals can infer the future state of the Shanghai Stock Index and give investors some advice and reference to help them made wise decisions. However, like many other auto-regressive models, fuzzy random autoregression model has some disadvantages, for example, it can perform well only in a placid market situation. On the contrary, when the market fluctuates fiercely because of some policy or man-behave influences, the model cannot get the meaningful results.

In this study, we built a fuzzy random autoregression model by employing expectations and variances of fuzzy random variables being used here to construct the confidence intervals of fuzzy random data. The proposed regressive model can be handled by solving a series of linear programming problems. The fuzzy random autoregression model can be used to forecast the confidence interval of future data in a time series data. In this study, we introduce the application of the fuzzy random

autoregression model in the stock market forecasting, it can also be applied in other time series forecasting such as temperature of an area or wave height of the ocean.

References

1. Yabuuchi, Y., Watada, J., Toyoura, Y.: Fuzzy AR model of stock price. Scientiae Mathematicae Japonicae 60(2), 303–310 (2004)
2. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
3. Ozawa, K., Watanabe, T., Kanke, M.: Forecasting Fuzzy Time Series with Fuzzy AR Model. In: Proceedings, the 20th International Conference on Computers & Industrial Engineering, pp. 105–108 (1996)
4. Kwakernaak, H.: Fuzzy random variables-I. Definitions and Theorems. Information Sciences 15(1), 1–29 (1978)
5. Hop, N.V.: Fuzzy stochastic goal programming problems. European Journal of Operational Research 176(1), 77–86 (2007)
6. Wang, G., Qiao, Z.: Linear programming with fuzzy random variable coefficients. Fuzzy Sets and Systems 57(3), 295–311 (1993)
7. Wang, S., Watada, J.: Reliability optimization of a series-parallel system withfuzzy random lifetimes. International Journal of Innovative Computing, Information & Control 5(6), 1547–1558 (2009)
8. Nahmias, S.: Fuzzy variables. Fuzzy Sets Syst. 1(2), 97–111 (1978)
9. Liu, B., Liu, Y.K.: Expected value of fuzzy variable and fuzzy expected value models. IEEE Trans. Fuzzy Syst. 10(4), 445–450 (2002)
10. Liu, Y.K., Liu, B.: Fuzzy random variable: A scalar expected value operator. Fuzzy Optim. Decision Making 2(2), 143–160 (2003)
11. Watada, J., Pedrycz, W.: A fuzzy regression approach to acquisition of linguistic rules. In: Pedrycz, W., Skowron, A., Kreinovich, V. (eds.) Handbook on Granular Comutation, ch. 32, pp. 719–740. John Wiley & Sons Ltd., Chichester (2008)

Building Linguistic Random Regression Model and Its Application

Sha Li, Shinya Imai, and Junzo Watada

Abstract. The objective of this paper is to build a model for the linguist random regression model as a vehicle to solve linguistic assessment given by experts. The difficulty in the direct measurement of certain characteristics makes their estimation highly impressive and this situation results in the use of fuzzy sets. In this sense, the linguistic treatment of assessments becomes essential when fully reflecting the subjectivity of the judgment process. When we know the attributes assessment, the linguistic regression model get the total assessment.

Keywords: Linguistic expression, fuzzy random variable, expected value, fuzzy regression model, variance, confidence interval.

1 Introduction

That we all know, experts with much professional experiences are capable of making assessment using their intuition and experiences. In the assessment process, the linguistic words are always used to do the evaluation. On this condition, we cannot get an accurate evaluation of the object. To cope with linguistic variables, we define processes of vocabulary translation and vocabulary matching which convert linguistic expressions into membership functions defined in the unit interval, and vice versa. Fuzzy random regression analysis [1], [5], is employed to deal with mapping and assessment process of experts which are realized from linguistic variables of features and characteristics of an objective into the linguistic expression articulating the total assessment.

The first fuzzy linear regression model, which treats fuzzy data instead of statistical data, was proposed by Tanaka in 1982. And in 1986, Watada and Tanaka

Sha Li · Shinya Imai · Junzo Watada
Graduate School of Information, Production and Systems, Waseda University,
2-7 Hibikino, Wakamatsu, Kitakyushu 808-0135 Japan
e-mail: lishacmos@gmail.com, shinimai76@gmail.com,
junzow@osb.att.ne.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 165–174.
springerlink.com

presented possibilistic regression based on the concept of possibility measure. Then in 2001 Chang discussed a fuzzy least-squares regression by using weighted fuzzy-arithmetic and the least-squares fitting criterion. In 2003, Y-K Liu and B.Liu presented a method to calculate the fuzzy random variables. In 2008 Watada presented that fuzzy regression approach to acquisition of linguistic rules. And in 2009 Watada proposed a new confidence-Interval-Based Fuzzy random regression models which can be used to solve the fuzzy random regression model.

In this paper all fuzzy sets are assumed normal triangular fuzzy numbers as (a, b, c), where 'b' denotes 'a' central value and 'a' and 'b' a left and right spreads, respectively. The paper consists of five sections. Section 2 introduces the definition of fuzzy random variables. Section 3 formulates the fuzzy linguistic random regression model with confidence interval and provides its solution method. Then in section 4, a simple application is given as an example. Then finally, Section 5 summarizes the conclusions.

2 Fuzzy Random Variable

Note that credibility measure is an average of the possibility and the necessity measure, i.e., Cr{·} = (Pos{·}+Nec{·})/2, and it is a self-dual set function, i.e., Cr{A} = 1- Cr{A^c} for any A in P(Γ). The motivation behind the introduction of the credibility measure is to develop a certain measure, which is a sound aggregate of the two extreme cases, such as the possibility (which expresses a level of overlap and is highly optimistic in this sense) and necessity (which articulates a degree of inclusion and is pessimistic in its nature). Based on credibility measure, the expected value of a fuzzy variable is presented as follows:

Definition 1. (Liu *et al.* [6]): Let Y be a fuzzy variable. The expected value of Y is defined as

$$E[Y] = \int_0^\infty Cr\{Y \geq r\}dr \int_{-\infty}^0 Cr\{Y \leq r\}\,dr \tag{1}$$

provided that at least two integrals are finite.

Example 1. The triangular fuzzy variable $\varepsilon = (a,b,c)$ has an expected value $E(\varepsilon) = \dfrac{a+2b+c}{4}$.

Definition 2. (Liu *et al.* [6]): Let ε be a fuzzy random variable with expected value e. Then, the variance of ε is defined by $V[\varepsilon] = E[(\varepsilon - e)^2]$.

Example 2. Let $\varepsilon = (a,b,c)$ be a symmetric triangular fuzzy variable, i.e., $b-a = c-b$. Then its variance is $V[\varepsilon] = \dfrac{(c-a)^2}{24}$.

3 Linguistic Fuzzy Random Regression Model

3.1 Linguistic Random Variables and Vocabulary Matching

In making assessments regarding some objects, we use multi-attribute evaluation. The difficulty in the direct measurement of certain characteristics makes their estimation highly impressive and this situation results in the use of fuzzy values and linguistic values. Often, experts use a linguistic word to judge an object from various features and characteristics. And the whole process is in linguistic way. For instance, although it is possible to measure numerical value, it is difficult to analytically interpret the obtained numerical value in terms of possible influence. This result might have impacted on further decision making. As Table 1 shows and evaluation of attribute k given by expert i.

Table 1 Linguistic values of each object w given by experts

sample	Attribute k Value			Total assessment
	1	$\cdots$	K	Y
1	(L_{11}, p_{11})	$\cdots$	(L_{K1}, p_{K1})	(T_1, p_1)
2	(L_{12}, p_{12})	$\cdots$	(L_{K2}, p_{K2})	(T_2, p_2)
3	(good,0.2)	$\cdots$	(very good,0.1)	(good,0.1)
$\vdots$	$\vdots$		$\vdots$	$\vdots$
N	$L_{N1}, p_{N1})$	$\cdots$	(L_{KN}, P_{KN})	(T_N, P_N)

In this study we built a model based on the relationship between the assessments given for different attributes and the overall assessment of the object totally. Watada *et al.*[1] propose fuzzy random regression model with confidence interval to deal with situations under hybrid uncertainty. The data given by experts are shown in Table 1 such as *"good," "bad," "extremely bad,"* as fuzzy random numbers.

An event has its population including the finite or infinite number of samples with probability. Generally such probability is not known clearly. We employ it by the linguistic assessment result percentage. Such as, 50 experts evaluate the object good, and 50 percentage evaluate the object very good, then the probability is 0.5, 0.5 respectively.

Then, we translate attributes from linguistic values *Li* into fuzzy grades X_L making use of triangular membership functions:

$$X_L \equiv (a, b, c) \tag{2}$$

where X_L denotes the central value of the fuzzy event, b, c are the left-side bound and right-side bound, respectively.

The estimation of the total assessment is written by the following fuzzy assessment function:

$$Y_i = f(X_{L1}, X_{L2}, \cdots, X_{LK}) \tag{3}$$

where *i=1, 2... N*, is the number of experts, *k* is the attributes of the object. Then the X_L is obtained from the vocabulary matching. Figure 2 is the dictionary of descriptive adjectives. From this dictionary we can convert the linguistic words to fuzzy variable random numbers. As according to the figure the linguistic word "*good*" = (2.5, 4.0, 5.5).

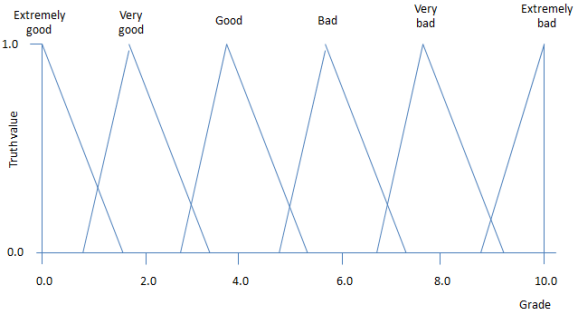


Fig. 1 Example figure for demonstration

3.2 Regression Model

Then all the linguistic data have been converted to fuzzy random variable data. Then we need to build a fuzzy regression model for fuzzy random data, which is based on the possibilistic linear model.

Fuzzy Random Regression Model with Confidence Interval: Table 2 is the format of data that come from linguistic words, where input data X_{ik} and output data Y_i, for all *i=1,2,...* , *n* and *k=1,2,...,K*. They are all fuzzy random variables, which defined as:

$$\begin{aligned} Y_i &= \bigcup_{t=1}^{M_{Yi}} \{(Y_i^t, Y_i^t, Y_i^t), p_i^t\}, \\ X_{ik} &= \bigcup_{t=1}^{M_{Xik}} \left\{ \left(X_{ik}^t, X_{ik}^{t,l}, X_{ik}^{t,r}\right), q_{ik}^t \right\}, \end{aligned} \tag{4}$$

respectively. This means that all values are given as fuzzy numbers with probabilities, where fuzzy variables (Y_i^t, Y_i^t, Y_i^t) and $\left(X_{ik}^t, X_{ik}^{t,l}, X_{ik}^{t,r}\right)$ are associated with probability p_i^t and q_{ik}^t, for $i = 1,2,\cdots,N$, $k = 1,2,\cdots,K$ and $t = 1,2,\cdots, M_{Yi}$ and M_{Xik} respectively.

Let us denote fuzzy linear regression model with fuzzy coefficients $\bar{A}_1,\cdots,\bar{A}_k$ as follows:

$$\bar{Y}_i = \bar{A}_1 X_{i1} + ... + \bar{A}_K X_{iK}, \tag{5}$$

And then we need to determine the optimal fuzzy parameters$\bar{A}_i$. Two optimization criteria are considered. One concerns the fitness of the fuzzy regression model, *h*.

The other one deals with fuzziness captured by the fuzzy regression model, S. Let us elaborate on the detailed formulation of these criteria.

As $\bar{A}_i$ is defined by $\bar{A}_i$= *(a, b, c)* which under the conditions that *w=1,2,. . . ,n,* where $h(w)$ appropriate linguistic words are selected by matching some value to a word in the linguistic dictionary $h(S)$ indicates the fitness of the estimated value with respect to a sample *w*and*h*0 denotes the fitness standard, while $a_i, \bar{a}_i$are defined as $\bar{A}_i^{ho} = [a_i, \bar{a}_i]$,where $\bar{Y}_i$denotes an estimate of the output and $\bar{A}_K = \left(\bar{A}_k^l + \bar{A}_k^r/2, \bar{A}_k^l, \bar{A}_k^r\right)_T$ are triangular fuzzy coefficients when triangular fuzzy random data X_{ik} are given for $i = 1,2,\cdots,N, k = 1,2,\cdots,K$, as shown in Table 2.

Table 2 Linguistic values of each object w given by experts

Sample	Inputs				Output
i	X_1	X_2	$\cdots$	X_K	Y
1	X_{11}	X_{12}	$\cdots$	X_{1K}	$Y1$
2	X_{21}	X_{22}	$\cdots$	X_{2K}	$Y2$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
N	X_{N1}	X_{N2}	$\cdots$	X_{NK}	YN

Then when we know the input and the output we can determine the fuzzy random linear model so that the model includes all given fuzzy random outputs. There the regression model with fuzzy random data can be built as the follow:

$$\left.\begin{array}{l} \min\limits_{\bar{A}} J(\bar{A}) = \sum\limits_{k=1}^{K} (\bar{A}_k^r - \bar{A}_k^l) \\ \text{subject to } \bar{A}_k^r \geq \bar{A}_k^l, \\ \bar{Y}_i = \bar{A}_1 X_{i1} + \ldots + \bar{A}_K X_{iK} \supset_{FR} Y_i, \end{array}\right\} \tag{6}$$

where $k = 1,2,\cdots,K$

Here, the fuzzy random inclusion relation $\supset_{FR}$ is very important to the model (6), which can be defined in various ways. Watada and Wang [8] used expectation based inclusion, and converted the fuzzy random regression model (6) to the following expected value regression model which corresponds to the conventional fuzzy regression model:

$$\left.\begin{array}{l} \min\limits_{\bar{A}} J(\bar{A}) = \sum\limits_{k=1}^{K} (\bar{A}_k^r - \bar{A}_k^l) \\ \text{subject to } \bar{A}_k^r \geq \bar{A}_k^l, \\ \bar{Y}_i = \sum_{k=1}^{K} \bar{A}_k E(X_{ik}) \supset_h E(Y_i), \end{array}\right\} \tag{7}$$

where $k = 1,2,\cdots,K$, and the $\supset_h$ denotes the fuzzy inclusion relation realized at level h.

In this study, we employ the confidence-interval based inclusion, which combines the expectation and variance of fuzzy random variables and the fuzzy

inclusion relation satisfied at level h, to deal with the model (7). There are also some other ways to define the fuzzy random inclusion relation $\supset_h$, which will yield more complicated fuzzy random regression models. For instance, in order to retain more complete information of the fuzzy random data, we can use the fuzzy inclusion relation directly for the product between a fuzzy parameter and a fuzzy value at some probability level. However, such calculation could be difficult since the product of two triangular fuzzy numbers does not retain the same triangular shape of the resulting membership function. Given this, the solution to the problem will rely on some heuristics method as proposed by Watada and Pedrycz [2].

First we consider the one sigma confidence interval of each fuzzy random variable, and it is expressed as follows:

$$\begin{aligned} I(e_{X_{ik}}, \sigma_{X_{ik}}) &= [e_{X_{ik}} - \sigma_{X_{ik}}, e_{X_{ik}} + \sigma_{X_{ik}}] \\ I(e_{Y_i}, \sigma_{Y_i}) &= [e_{Y_i} - \sigma_{Y_i}, e_{Y_i} + \sigma_{Y_i}] \end{aligned} \tag{8}$$

Then the new confidence-interval-based fuzzy random regression mode is built as follows:

$$\left.\begin{aligned} \min_{\bar{A}} J(\bar{A}) &= \sum_{k=1}^{K} (\bar{A}_k^r - \bar{A}_k^l) \\ \text{subject to } \bar{A}_k^r &\geq \bar{A}_k^l, \\ \bar{A}_i &= \sum_{k=1}^{K} \bar{A}_k I[e_{X_{ik}}, \sigma_{X_{ik}}] \subset_h I[e_{Y_i}, \sigma_{Y_i}] \end{aligned}\right\} \tag{9}$$

For *i=1, 2, …,N, k=1,2,…,K*

As calculation is very complicated, a heuristic method is used to solve this model. This heuristic algorithm was expanded from the one proposed by Watada in 1985[7]. At a α-level set of the fuzzy degree of a structural attribute X_{ik}at ho is denoted by (9). There will be 5 steps to calculate:

Step 1: Let the trial count n=1 and let use consider, the termination count is set to N, and using e_{ik} of attributes k=1,2,…,K, for each sample i=1,2,…,N, determine both the bounds $a_k^{(n)}, \bar{a}_k^{(n)}$ of $\bar{A}_k^{(n)}$ by solving the linear programming task (8).

Step 2: Determine $(\bar{A}_K^{(n+1)} gI[e_{X_{ik}}, \sigma_{X_{ik}}])_{h_0}$ according to the table 3, by assigning the signs of $a_k^{(n)}$ and $\bar{a}_k^{(n)}$ to $a_k^{(n+1)}$ and $\bar{a}_k^{(n+1)}$, for k=1,2,…K. Determine $a_k^{(n+1)}$ and $\bar{a}_k^{(n+1)}$ for k=1,2,…,K by the linear programming to minimize the fuzziness $J(\bar{A})$ under the constraints in the programming problem (10) according to the conditions explained in Table 3.

Step 3: If $a_k^{(n+1)} \cdot a_k^{(n+1)} \geq 0$ and $\bar{a}_k^{(n+1)} \cdot \bar{a}_k^{(n)} \geq 0$(*k=1, 2,…, K*), then go to step 5. Otherwise, let n=n+1, and go to step 4.

Step 4: If the trial count N has not exceeded the given termination count N, then go to step 2. Otherwise, go to step 5.

Step 5: Then solve the LP problem and calculate the $a_k^{(n)}$, $\bar{a}_k^{(n)}$, $a_k^{(n+1)}$, $\bar{a}_k^{(n+1)}$ and go to step 6.

Step 6: Output the solution and terminate the algorithm.

4 A Simple Application

Table 3 Output and Inputs

Sample	Input			Output
i	X_1	X_2	X_3	Y
1	(Extremely good), 0.5; (very good), 0.5;	(very good), 0.6; (good), 0.4 ;	(bad), 0.5; (very bad), 0.5 ;	(popular), 0.8; (normal), 0.2;
2	(very good), 0.75; (good), 0.25;	(good), 0.6; (very good), 0.4;	(good), 0.8 ; (bad), 0.2;	(popular), 0.6; (normal), 0.4;
3	(good), 0.6 ; ((bad), 0.4;	(very good), 0.8; (good), 0.2;	(extremely good), 0.6; (good), 0.4;	(popular), 0.75 ; (normal), 0.25 ;
4	(good), 0.2; (very bad), 0.8;	(bad), 0.4 ; (very bad), 0.6;	(bad), 0.6; (extremely bad), 0.4;	(bad), 0.6 ; (normal), 0.4;

20 questionnaires were sent to 20 people to evaluate four hotels in china. Then returned answers are collected and results are partly in Table 3. The sample is 4 hotels and X_1, X_2, X_3, is the attributes that we use to evaluate the hotels, they are accommodation state, environment around state and facilities state, respectively. Then, according to Figure 1, we convert linguistic words in Table 3 to fuzzy random variables as shown in Tables 4 and 5. The probability is the percent number of the assessment. Such as, when the probability is 0.5, where the answers of the questionnaire there are 50 percent of people give this evaluation.

Table 4 Fuzzy sets of inputs and output

sample	Accommodation state /X_1	Environment around state/X_2	Facilities state/X_3	Total Assessment/Y
1	(0.0,0.0,1.5), 0.5	(0.5,2.0,3.5), 0.6	(4.5,6.0,7.5), 0.5	(0.0,0.0,4.0), 0.8
	(0.5,2.0,3.5), 0.5	(2.5,4.0,5.5), 0.4	(6.5,8.0,9.5), 0.5	(3.0,5.0,7.0), 0.2
2	(0.5,2.0,3.5), 0.75	(2.5,4.0,5.5), 0.6	(2.5,4.0,5.5), 0.8	(0.0,0.0,4.0), 0.6
	(2.5,4.0,5.5), 0.25	(0.5,2.0,3.5), 0.4	(4.5,6.0,7.5), 0.2	(3.0,5.0,7.0), 0.4
3	(2.5,4.0,5.5), 0.6	(0.5,2.0,3.5), 0.8	(0.0,0.0,1.5), 0.6	(0.0,0.0,4.0), 0.75
	(4.5,6.0,7.5), 0.4	(2.5,4.0,5.5), 0.2	(2.5,4.0,5.5), 0.4	(3.0,5.0,7.0), 0.25
4	(2.5,4.0,5.5), 0.2	(4.5,6.0,7.5), 0.4	(4.5,6.0,7.5), 0.6	(6.0,6.0,10.0), 0.6
	(6.5,8.0,9.5), 0.8	(6.5,8.0,9.5), 0.6	(8.5,8.5,10.0), 0.4	(3.0,5.0,7.0), 0.4

The fuzzy regression model with confidence interval for the given data reads as follow:

$$\bar{Y}_i = \bar{A}_1\,[e_{Xi1}, \sigma_{Xi1}] + \bar{A}_2\,[e_{Xi2}, \sigma_{Xi2}] + \bar{A}_3\,[e_{Xi3}, \sigma_{Xi3}]$$

where $[e_{Xi1}, \sigma_{Xi1}]$, for $k = 1,2,3$, from the model (9), and assuming $(\bar{A}_K)_{h0} = [\bar{A}_k^l, \bar{A}_k^r], k = 1,2,3$, the model can be built as :

$$\begin{aligned}
\min_{\bar{A}}\ & J(\bar{A}) = A_1^r - A_1^l + A_2^r - A_2^l + A_3^r - A_3^l \\
\text{subject to}\ & \bar{Y}_1 = [A_1^r, A_1^l] * I[e_{X_{11}}, \sigma_{X_{11}}] + [A_2^r, A_2^l] * I[e_{X_{12}}, \sigma_{X_{12}}] + [A_3^r, A_3^l] * I[e_{X_{13}}, \sigma_{X_{13}}] \\
& \bar{Y}_2 = [A_1^r, A_1^l] * I[e_{X_{21}}, \sigma_{X_{21}}] + [A_2^r, A_2^l] * I[e_{X_{22}}, \sigma_{X_{22}}] + [A_3^r, A_3^l] * I[e_{X_{23}}, \sigma_{X_{23}}] \\
& \bar{Y}_3 = [A_1^r, A_1^l] * I[e_{X_{31}}, \sigma_{X_{31}}] + [A_2^r, A_2^l] * I[e_{X_{32}}, \sigma_{X_{32}}] + [A_3^r, A_3^l] * I[e_{X_{33}}, \sigma_{X_{33}}] \\
& \bar{Y}_4 = [A_1^r, A_1^l] * I[e_{X_{41}}, \sigma_{X_{41}}] + [A_2^r, A_2^l] * I[e_{X_{42}}, \sigma_{X_{42}}] + [A_3^r, A_3^l] * I[e_{X_{43}}, \sigma_{X_{43}}]
\end{aligned} \tag{10}$$

First, we need to calculate the $I[e_{Xik}, \sigma_{Xik}]$ and $I[eY_{i,}\sigma Y_{i,}]$ for the input and output. By using the calculation in Examples 1 and 2, we get the data as shown in Tables 6, and 7. Then the confidence intervals for the input data and output data can be calculated in the form (8). They are listed in Table 8.

Table 5 Expectation and Standard deviation of the Data

i	$(e_{x_{i1}}, \sigma_{x_{i1}})$	$(e_{x_{i2}}, \sigma_{x_{i2}})$	$(e_{x_{i3}}, \sigma_{x_{i3}})$	(e_{Y_i}, σ_{Y_i})
1	(1.19,0.234)	(2.8,0.375)	(7.0,0.375)	(1.8,0.667)
2	(2.5,0.375)	(3.2,0.375)	(4.4,0.375)	(2.6,0.667)
3	(4.8,0.375)	(2.4,0.375)	(1.83,0.206)	(2.0,0.667)
4	(7.2,0.375)	(7.2,0.375)	(7.15,0.263)	(6.2,0.667)

Table 6 Confidence intervals of the Data

i	$I(e_{x_{i1}}, \sigma_{x_{i1}})$	$I(e_{x_{i2}}, \sigma_{x_{i2}})$	$I(e_{x_{i3}}, \sigma_{x_{i3}})$	$I(e_{Y_i}, \sigma_{Y_i})$
1	(0.956,1.424)	(2.425,3.175)	(6.625,7.375)	(1.033,2.367)
2	(2.125,2.875)	(2.825,3.575)	(4.025,4.775)	(1.733,3.067)
3	(4.425,5.175)	(2.025,2.775)	(1.642,2.036)	(1.333,2.667)
4	(6.825,7.575)	(6.825,7.575)	(6.887,7.413)	(6.733,6.867)

Then, the model (10) can be rewritten as:

$$\begin{aligned}
\min_{\bar{A}}\ & J(\bar{A}) = A_1^r - A_1^l + A_2^r - A_2^l + A_3^r - A_3^l \\
\text{subject to}\ & \bar{Y}_1 = [A_1^r, A_1^l] * I[0.956, 1.424] + [A_2^r, A_2^l] * I[2.425, 3.175] \\
& \quad + [A_3^r, A_3^l] * I[6.625, 7.375] \supset I[1.033, 2.367] \\
& \bar{Y}_2 = [A_1^r, A_1^l] * I[2, 125, 2.875] + [A_2^r, A_2^l] * I[2.825, 3.575] \\
& \quad + [A_3^r, A_3^l] * I[4.025, 4.775] \supset I[1.733, 3.067] \\
& \bar{Y}_3 = [A_1^r, A_1^l] * I[4.425, 5.175] + [A_2^r, A_2^l] * I[2.025, 2, 775] \\
& \quad + [A_3^r, A_3^l] * I[1.642, 2.036] \supset I[1.333, 2.667] \\
& \bar{Y}_4 = [A_1^r, A_1^l] * I[6.825, 7.575] + [A_2^r, A_2^l] * I[6.825, 7.575] \\
& \quad + [A_3^r, A_3^l] * I[6.887, 7.413] \supset I[6.733, 6.867]
\end{aligned} \tag{11}$$

According the algorithm we solve the linear problem with lingo, we get the answer:

$$\begin{aligned}
& \min_{\bar{A}} J(\bar{A}) = A_1^r - A_1^l + A_2^r - A_2^l + A_3^r - A_3^l \\
& = 0.18 - 0.13 + 0.37 - 0.37 + 0.348 - 0.0 \\
& = 0.398
\end{aligned} \tag{12}$$

And coefficient is following,

$$\begin{aligned} \bar{A}_1^r &= 0.18, \ \bar{A}_1^l = 0.13, \\ \bar{A}_2^r &= 0.37, \ \bar{A}_1^l = 0.37, \\ \bar{A}_3^r &= 0.00, \ \bar{A}_1^l = 0.348 \end{aligned} \tag{13}$$

Thus, the fuzzy random regression model with confidence interval is given in the form:

$$\begin{aligned} \bar{Y}_i = &[0.13, 0.178] \times I[e(X_{ik}), \sigma(X_{ik})] + [0.37.0.37] \times I[e(X_{ik}), \sigma(X_{ik})] \\ &+[0.0, 0.348] \times I[e(X_{ik}), \sigma(X_{ik})] \end{aligned} \tag{14}$$

Then we input the original data to test the model's accurate, we can get the following comparison results in Table 8. From the results, we can see this model is work good, that all the result is all concluded in the interval calculated by the model. We can see all the total evaluation can be predicted by this model. And for sample 2 and 3, it gives an accurate result. Then we need to match the numbers to the linguistic total assessment words. For this, we draw a fuzzy grade picture for the total assessment estimated by the model in Figure 2. We can see the area covered is almost same as we get from the people. The A, B, C, D is the samples 1,2,3,4. From the area of the triangle, we can see approximately probabilities of the total assessment.

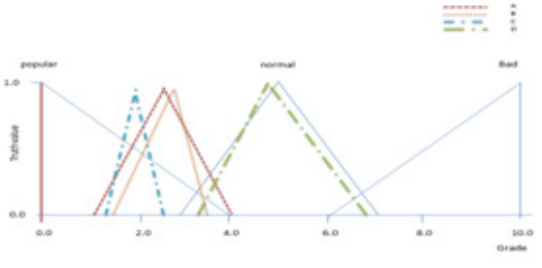


Fig. 2 Example figure for demonstration

Table 7 Comparison of the expert evaluation and the model

sample	Experts value	Estimated fuzzy grade	Matched word
1	(1.033,2.367)	(1.021,4.012)	Popular
2	(1.733,3.067)	(1.321,3.511)	Popular
3	(1.333,2.667)	(1.324,2.671)	popular
4	(6.733,6.867)	(3.412,6.876)	normal

5 Conclusions

As we all know, human experts are always involved in decision-making process. However, the judgment experience and knowledge of experts are unique to each other. Better understanding of this judgment knowledge, sometimes, we need to convert it to numerical values which can give people more ocular way to experience the whole assessment process. And at the same time, there is always linguistic assessment of an object from various attributes. Then it is difficult to get a total assessment when we have linguistic data. In this paper, our model is built to solve this kind of problem. However, there is still a problem we need to solve. It is the problem of the matching process. Finally we get the confidence interval of the sample. However, we cannot get the exact probabilities that attached with the linguistic word. So later, we will try to solve this problem.

References

1. Imai, S., Wang, S., Watada, J.: Regression Model Based on Fuzzy Random Variables. In: KES 2008, Part III. LNCS (LNAI), vol. 5179, pp. 127–135. Springer, Heidelberg (2008)
2. Watada, J., Pedrycz, W.: A fuzzy regression approach to acquisition of linguistic rules. In: Pedrycz, W. (ed.) Handbook on Granular Commutation, ch.32, pp. 719–730 (2008)
3. Kwakernaak, H.: Fuzzy random variables “I. Denitions and theorems”. Information Sciences 15(1), 1–29 (1978)
4. Kwakernaak, H.: Fuzzy random variables “II. Algorithm and examples”. Information Sciences 17(3), 253–278 (1979)
5. Liu, B.: Fuzzy random chance-constrained programming. IEEE Transactions on Fuzzy Systems 9(5), 713–720 (2001)
6. Liu, B., Liu, Y.-K.: Expected value of fuzzy variable and fuzzy expected value models. IEEE Transactions on Fuzzy Systems 10(4), 445–450 (2002)
7. Liu, Y.-K., Liu, B.: Fuzzy random variable: A scalar expected value operator. Fuzzy Optimization and Decision Making 2(2), 143–160 (2003)
8. Watada, J., Wand, S., Pedrycz, W.: Building confidence-interval-based fuzzy random regression model. IEEE Transactions on Fuzzy Systems 11(6), 1273–1283 (2009)

Classifying Attributes with Game-Theoretic Rough Sets

Nouman Azam and JingTao Yao

Abstract. Attributes may be categorized as core, reduct or non-reduct attributes when rough set theory is utilized for classification. These attribute types play different roles in feature selection algorithms. We introduce a game-theoretic rough set based method that formulates the classification of an attribute as a decision problem within a game. In particular, multiple measures representing importance levels for an attribute are incorporated into a unified framework to obtain an effective attribute classification mechanism. Demonstrative example suggests that the method may be efficient in classifying different types of attributes.

1 Introduction

Rough set theory introduces the concept of reducts which can be used in categorizing attributes as core, reduct or non-reduct attributes, respectively [1]. Reduct is a minimal subset of attributes that has the same descriptive or classification ability as the entire attribute set [2, 3]. There may exist more than one reduct in a given dataset.

The notion of multiple reducts leads to three different types of attributes. An attribute appearing in every reduct is called a core attribute; If it appears in at least one reduct, it is called a reduct attribute; If it does not appear in any of the reduct, it is called a non-reduct attribute [1]. Wei et al. [4] and Zhang et al. [5] referred to these attribute types as absolutely necessary, relatively necessary or absolutely unnecessary. Hu et al. [6] used the terms strongly relevant, weakly relevant or completely irrelevant. The core attributes are essential in making correct classification for a data set and therefore they should be retained. The non-reduct attributes are redundant ones for classification purpose and should be eliminated. The reduct attributes may be considered as attributes having certain level of importance [7].

Nouman Azam · JingTao Yao
Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
e-mail: {azam200n,jtyao}@cs.uregina.ca

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 175–184.
springerlink.com

The starting point of classifying attributes in rough sets is to find most, if not all, reducts. There exists many methods for obtaining multiple reducts [8, 9, 10, 11]. These methods commonly involve an iterative process where each iteration involves a sub-iterative process for searching a single reduct. A particular attribute may be processed and evaluated multiple times during different iterations of these methods for determining a right search direction. Such multiple processing of attributes may affect the performance of these methods for finding multiple reducts. As a matter of fact, finding all reducts has been established as an NP-hard problem, requiring extensive computations [12, 13]. We try to find additional mechanism for classifying attributes within the rough sets. Particularly, we look for a method of attribute classification that processes each attribute once therefore avoiding extensive computations. A game-theoretic rough sets (GTRS) based method is introduced for such a purpose.

The GTRS model is a recent extension to rough set theory that allows for simultaneous consideration of multiple criteria in order to analyze and make intelligent decisions [14, 15, 16]. The model interprets effective decision making among multiple decisive factors as a competitive or cooperative process observed with game-theoretic analysis. We utilize the model in formulating a method that interprets the classification of a feature as a decision problem within a game involving multiple measures. Each measure analyzes an attribute for its importance using suitable payoff functions and may choose from two possible actions for an attribute namely, accept (or select) and reject. The actions of respective measures are utilized in classifying an attribute into one of the three possible attribute sets.

2 Game-Theoretic Rough Sets

The game-theoretic rough set model was proposed by Yao and Herbert as an extension to probabilistic rough sets [14]. A mechanism was proposed to decrease the size of the boundary region by introducing an alternative approach for deriving the probabilistic rough set regions. We will briefly review the model in this section.

The GTRS model utilizes ideas from game theory in analyzing decision problems [17]. Game theory has been a core subject in decision sciences being utilized in many areas of computer science such as machine learning, cryptography and networking [18, 19, 20, 21]. A single game may be formally expressed as $G = \{P, S, F\}$, where $P = \{p_1, p_2, ..., p_n\}$ represents the player set, $S = \{a_1, a_2, ..., a_m\}$ the available strategy or action set for each player, and $F = \{u_1, u_2, ..., u_m\}$ the respective payoff functions for each action.

In GTRS formulation, a player set P contains multiple critical factors or criteria that highlight particular aspects of interest in a give decision problem. The selection of these factors significantly depends on the goals we seek to attain. For instance, if we wish to increase the overall classification ability of the system then the factors reflecting the classification capabilities, i.e. precision, recall and accuracy may be considered as possible players for a game.

The GTRS formulates strategies for each measure to support their effective participation in a game. Strategies may represent various alterations to variables or parameters involved in a decision problem that affect decision making. For example, changing the parameters corresponding to loss functions in decision theoretic rough sets may lead to different decisions for classifying objects [22]. The strategies which are realized as possible changes to variables may produce different levels of responses when considered with different criteria. This suggests that a particular criterion may experience a certain level of performance gain by selecting an appropriate variable with its associated change. A gain observed in a such selection may be considered as a payoff or utility for a criterion.

Table 1 Action scenario for a measure

Action	Description	Method	Desired conditions	Outcomes or Influences
	What the action does?	How to carry out the action?	When the action is desired?	Possible effects on problem
a_j	....	....	....	....
....	....	....	....	....

A player participates in a game by selecting appropriate actions in different situations he may encounter during the play. Particularly, each action is desired in a specific situation of a game. The GTRS constructs action scenarios for analyzing the actions of a player under specified situations. Table 1 presents the general structure of an action scenario where each row represents an action with its associated details.

The objective of GTRS is to assist in intelligent decision making. Payoff tables are constructed for such a purpose that lists all possible strategies for each considered criterion along with their respective payoffs. The payoff tables are examined with equilibrium analysis to determine an optimal solution. This means that we try to find a solution where each measure has maximized its payoff given their opponents chosen actions.

3 A GTRS Based Method for Classifying Attributes

The notion of reducts in rough sets can be used to classify objects without deteriorating the classification quality. Let $R_1, R_2, ..., R_k$ be all possible reducts in an information system. An attribute A is identified as, a core attribute, if $A \in \bigcap_{i=1}^{k} R_i$, a reduct attribute, if $A \in \bigcup_{i=1}^{k} R_i$, and a non-reduct attribute, if $A \notin \bigcup_{i=1}^{k} R_i$ [4]. We will use the sets $CORE$, $REDUCT$ and $NREDUCT$ to represent core, reduct and non-reduct attribute sets as suggested by Yao et al. [1].

We may need to compute all reducts in order to classify and obtain the three types of attributes. Existing methods for computing all reducts involve extensive computations. We try to simplify the classification of attributes by introducing a classification mechanism which may be utilized in making reasonable predictions for classifying attributes. It was suggested in previous research that different measures of an attribute importance may lead to different reducts [23]. This means that multiple reducts may be obtained by making use of multiple measures. Once multiple reducts are obtained, we may utilize them in classifying attributes. We utilized a similar idea in formulating a method that incorporates multiple measures of an attribute importance in classifying an attribute.

3.1 Formulating a GTRS Based Method

We need to formulate problems in order to analyze them with the GTRS model. We start by identifying the main components of the model briefly discussed in Section 2, i.e. players in a game that represent different criteria in analyzing a problem, the available strategies or actions for each player and the payoff functions for respective actions.

The Players: We selected measures of an attribute importance as players in a game. The purpose of the game is to obtain an effective decision observed with cooperation among multiple measures for classifying a particular attribute. We consider two measures in this study, represented as m_1 and m_2. Each measure will determine an attribute importance independently while jointly they will cooperate to reach an optimal decision for classifying an attribute.

The Strategies (of Players): Each measure is expected to have a set of possible strategies or actions. The available actions for a player reflect various allowable moves that may be utilized during a game. We wish to formulate actions for measures that reflect their decisions on an attribute being accepted (selected) or rejected for a reduct. In particular, if a measure returns a high value for an attribute we may decide to select while in case of a lower value we may choose to reject. The actions of selecting and rejecting an attribute are represented as a_s and a_r, respectively.

3.2 Payoff Functions in Different Situations of a Game

A payoff function represents the motivation of a player towards an action. The payoff functions corresponding to different actions may be analyzed by considering individual actions of measures. We constructed action scenario given as Table 2 for this purpose. The first row represents an action of selecting an attribute. This action is desired when a particular measure m_i returns a higher value for an attribute A. The second row represents an action of rejecting an attribute. This action is desired when a measure returns a lower value for an attribute. This means that we may specify the payoff for measure m_i in selecting an attribute A as $m_i(A)$ while the payoff

for rejecting the same attribute as 1-$m_i(A)$. We will be using the notation $u_{i(j|k)}$ for representing a payoff function of player i, performing an action j, given an action k of his opponent.

Table 2 Action scenario for players

Action	Description	Desired Conditions	Outcome
a_s	Select	High value of $m_i(A)$	Feature is useful for reduct
a_r	Reject	High value of $1 - m_i(A)$	Feature is not useful for reduct

We now analyze the individual payoffs (shown in Table 2) in three possible situations of the game. In the first case, both measures choose to select an attribute. We wish to find payoffs cooperatively as both measures have consensus about an attribute. An average of the values for an attribute (calculated with respective measures) may be useful to represent the payoff functions of players in this case. This suggests that the measures will have high payoff only when both measures return high values. A second situation arises during the game when the measures have conflicting believes regarding an attribute. This means that one of the measures chooses the action a_s while the other a_r. As there is a conflict between the measures, payoffs are determined independently. Each measure receives either a payoff equals to $m_i(A)$ or 1-$m_i(A)$ depending on the chosen action of a_s or a_r, respectively. Finally, we may encounter a situation in which both measures choose to reject an attribute. An average of the values with respective measures (for rejecting an attribute) are used to represent the payoff functions of players.

According to discussion in above cases, the payoff functions for a particular measure m_i with its opponent represented as m_{-i}, are calculated as follows.

$$\begin{aligned} u_{m_i(a_s|a_s)} &= \tfrac{m_i(A)+m_{-i}(A)}{2}, & u_{m_i(a_s|a_r)} &= m_i(A), \\ u_{m_i(a_r|a_r)} &= \tfrac{(1-m_i(A))+(1-m_{-i}(A))}{2}, & u_{m_i(a_r|a_s)} &= 1 - m_i(A). \end{aligned} \tag{1}$$

3.3 Competition Analysis and Attribute Classification Algorithm

We finally express the game as competition between the measures. The payoff tables, which lists all possible actions with respective utilities are used for this purpose. Table 3 presents a payoff table with rows and columns representing the players m_1 and m_2, respectively. Each entry in the table represents a payoff pair $< u_{m_1(a_i|a_j)}, u_{m_2(a_j|a_i)} >$ corresponding to actions a_i of m_1 and a_j of m_2. A total of four payoff pairs are required in this case.

Table 3 The payoff table

		m_2	
		a_s	a_r
m_1	a_s	$u_{m_1(a_s\|a_s)}, u_{m_2(a_s\|a_s)}$	$u_{m_1(a_s\|a_r)}, u_{m_2(a_r\|a_s)}$
	a_r	$u_{m_1(a_r\|a_s)}, u_{m_2(a_s\|a_r)}$	$u_{m_1(a_r\|a_r)}, u_{m_2(a_r\|a_r)}$

We compute Nash equilibrium [17] within the payoff table for examining possible outcomes of the game. This intuitively suggests that none of the measures can be benefited by changing his strategy, given his opponents chosen actions. The pair $(u^*_{m_1(a_i|a_1)}, u^*_{m_2(a_j|a_2)})$ is an equilibrium if for any action a_k, where $k \neq i, j$, $u^*_{m_1(a_i|a_1)} \geq u_{m_1(a_k|a_1)}$ and $u^*_{m_2(a_j|a_2)} \geq u_{m_2(a_k|a_2)}$.

Algorithm 1 GTRS based method for classifying attributes

Input: A set of attributes denoted as At.
Measures m_1 and m_2 for evaluation of attribute importance.
Output: $CORE$, $REDUCT$ and $NREDUCT$ sets.

1: Initialize $CORE = \emptyset$, $REDUCT = \emptyset$, $NREDUCT = \emptyset$.
2: **for all** attributes $a \in At$ **do**
3: Calculate the values of a according to measures m_1 and m_2.
4: Generate a payoff table for attribute a (according to Equation 1).
5: Perform equilibrium analysis within the payoff tables and determine the selected actions of measures.
6: **if** both measures choose the action a_s **then**
7: Classify attribute a as a core attribute, i.e. $CORE = CORE \bigcup \{a\}$.
8: **end if**
9: **if** one measure chooses the action a_s while the other a_r. **then**
10: Classify attribute a as a reduct, i.e. $REDUCT = REDUCT \bigcup \{a\}$.
11: **end if**
12: **if** both measures choose the action a_r **then**
13: Classify attribute a as a non-reduct, i.e. $NREDUCT = NREDUCT \bigcup \{a\}$.
14: **end if**
15: **end for**

Algorithm 1 summarizes the proposed method. From line 2 of the algorithm, we note that each attribute is processed exactly once. Each attribute is analyzed with game-theoretic analysis for identifying its possible classification. The three if statements in algorithm determine the classification of an attribute.

4 A Demonstrative Example

We now demonstrate the application of the proposed method for classifying attributes with an example. The information table of Table 4 is considered for this purpose. By considering the attribute g as a decision attribute in the table, we may compute reducts with respect to attribute g. According to properties of reducts [23], we note that there are only two possible reducts in the table, i.e. $\{a,e\}$ and $\{c,e\}$. This means that these attribute sets contains minimal possible attributes that have the same classification ability as the entire attribute set. We can compute the three attribute sets previously discussed in Section 3 as, $CORE = \{e\}$, $REDUCT = \{a,c,e\}$, and $NREDUCT = \{b,d,f\}$.

Table 4 Information table

	a	b	c	d	e	f	g
o_1	0	0	0	1	0	0	0
o_2	0	1	2	0	0	0	0
o_3	2	1	3	0	0	0	0
o_4	3	2	2	1	1	0	1
o_5	1	1	3	0	1	0	1
o_6	2	2	1	1	2	0	1
o_7	3	2	2	1	2	0	2
o_8	1	1	4	0	2	0	2
o_9	1	1	3	0	2	0	2

In order to measure the importance of attributes in examining the proposed method, two measures were selected, i.e. conditional entropy and dependency [24]. The two measures may be defined for a particular attribute A with respect to decision attribute set D as follows,

$$E(D|A) = -\sum_{A_i \in A} P(A_i) \sum_{D_k \in D} P(D_k|A_i) log P(D_k|A_i) \tag{2}$$

$$\gamma_A(D) = \frac{\|POS_A(D)\|}{\|U\|} \tag{3}$$

Equation 2 representing the conditional entropy, may be understood as the uncertainty of decision variable D conditional upon variable A. In other words, it is the remaining entropy (or uncertainty) in variable D given the information in variable A. The measure of dependency shown as Equation 3 may be defined as the fraction of rough set based positive region (induced by attribute D) that may be identified with the information of attribute A. We will be using $m_1(A)$ and $m_2(A)$, for referring to $E(D|A))$ and $\gamma_A(D)$, respectively for simplicity. Furthermore, normalized values of attributes returned with the measures would be utilized.

Table 5 Payoff table for attribute e

		m_2	
		a_s	a_r
m_1	a_s	**1.0,1.0**	1.0, 0.0
	a_r	0.0, 1.0	0.0, 0.0

Let us first analyze the core attribute e with this method. Its payoff table may be determined by utilizing equations listed as Equation 1. Table 5 shows its payoff table. The cell with bold numbers represents a Nash equilibrium. The actions of players in equilibrium corresponds to a_s for both measures. We note that none of the measures can achieve a higher payoff, given the other measure's chosen actions. Changing the actions of measures from a_s to a_r will result in a payoff decrease from 1.0 to 0.0 for both measures. The actions of measures in this case mutually predicts the attribute e as a core attribute, therefore it may be included in the set $CORE$.

Table 6 Payoff table for attribute a

		m_2	
		a_s	a_r
m_1	a_s	0.54, 0.54	**0.68,0.60**
	a_r	0.32, 0.40	0.46, 0.46

Table 7 Payoff table for attribute c

		m_2	
		a_s	a_r
m_1	a_s	0.52, 0.52	0.43, 0.40
	a_r	**0.57,0.60**	0.48, 0.48

We now consider the reduct attributes which include a, c and e. Table 6 shows the payoff table for attribute a. The equilibrium in this case corresponds to actions a_s for m_1 and a_r for m_2. We note again that both measures have maximized their payoffs, given their opponents chosen actions. Changing the action of m_1 from a_s to a_r will result in a payoff decrease from 0.68 to 0.46, while changing the action of m_2 from a_r to a_s will result in decrease from 0.60 to 0.54. From actions of the players, attribute a is included in the set $REDUCT$. Similarly, for attribute c, the equilibrium corresponds to its inclusion in the set $REDUCT$. This is presented in Table 7. The attribute e which is both reduct and core attribute has already been examined in this section.

We finally consider the non-reduct attributes, i.e. b, d and f. The payoff pairs for attribute b corresponding to action pairs $< a_s, a_s >, < a_s, a_r >, < a_r, a_s >, < a_r, a_r >$, are calculated as $< 0.28, 0.28 >, < 0.35, 80 >, < 0.65, 20 >$ and $<$ **0.72,0.72** $>$, respectively. The pair $<$ **0.72,0.72** $>$ was found to be an equilibrium which corresponds to an action of rejecting for both measures. Hence, attribute b is included in

the set $NREDUCT$. Similarly, by examining the attributes d and f, we found that equilibrium analysis leads to their inclusion in the set $NREDUCT$.

By observing the sets $CORE$, $REDUCT$ and $NREDUCT$ in above example, we note that the GTRS based method correctly classifies the attributes. Therefore, the GTRS based method may be useful for classifying attributes in rough sets.

5 Conclusion

Rough set theory can be utilized to classify attributes with reducts. Existing approaches for finding reducts commonly involve an iterative process where each iteration contains a sub-iterative process for searching a single reduct. A particular attribute may be processed many times due to multiple iterations in these methods. We introduce a game-theoretic rough set (GTRS) based method which provides an attribute classification mechanism by processing each attribute only once. The model was previously suggested for probabilistic rough sets for obtaining probabilistic rough set regions through simultaneous consideration of multiple influential factors or criteria. Particularly, a method is introduced that incorporates multiple measures of an attribute importance for classifying individual attributes. A demonstrative example suggests that the proposed method may be useful in classification of attributes while ensuring that each attribute is processed exactly once.

The classification results achieved with this method may be further refined by investigating interesting measures of an attribute's importance that is specifically designed for the game-theoretic framework proposed in this article.

References

1. Yao, Y., Zhao, Y., Wang, J., Han, S.: A Model of User-oriented Reduct Construction for Machine Learning. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets VIII. LNCS, vol. 5084, pp. 332–351. Springer, Heidelberg (2008)
2. Yao, Y.Y., Zhao, Y.: Discernibility matrix simplification for constructing attribute reducts. Information Science 179(7), 867–882 (2009)
3. Yao, J.T., Zhang, M.: Feature Selection with Adjustable Criteria. In: Ślęzak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3641, pp. 204–213. Springer, Heidelberg (2005)
4. Wei, L., Li, H.-R., Zhang, W.-X.: Knowledge reduction based on the equivalence relations defined on attribute set and its power set. Information Science 177(15), 3178–3185 (2007)
5. Zhang, M., Xu, L.D., Zhang, W.-X., Li, H.-Z.: A rough set approach to knowledge reduction based on inclusion degree and evidence reasoning theory. Expert Systems 20(5), 298–304 (2003)
6. Hu, Q., Liu, J., Yu, D.: Mixed feature selection based on granulation and approximation. Knowledge Based Systems 21(4), 294–304 (2008)
7. Hu, X., Lin, T.Y., Han, J.: A new rough sets model based on database systems. Fundamenta Informaticae 59(2-3), 135–152 (2004)

8. Ishii, N., Morioka, Y., Bao, Y., Tanaka, H.: Control of Variables in Reducts - kNN Classification with Confidence. In: König, A., Dengel, A., Hinkelmann, K., Kise, K., Howlett, R.J., Jain, L.C. (eds.) KES 2011, Part IV. LNCS, vol. 6884, pp. 98–107. Springer, Heidelberg (2011)
9. Strąkowski, T., Rybiński, H.: A New Approach to Distributed Algorithms for Reduct Calculation. In: Peters, J.F., Skowron, A., Rybiński, H. (eds.) Transactions on Rough Sets IX. LNCS, vol. 5390, pp. 365–378. Springer, Heidelberg (2008)
10. Liu, H., Abraham, A., Yue, B.: Nature Inspired Multi-Swarm Heuristics for Multi-Knowledge Extraction. In: Koronacki, J., Raś, Z.W., Wierzchoń, S.T., Kacprzyk, J. (eds.) Advances in Machine Learning II. SCI, vol. 263, pp. 445–466. Springer, Heidelberg (2010)
11. Wu, Q., Bell, D.A.: Multi-Knowledge Extraction and Application. In: Wang, G., Liu, Q., Yao, Y., Skowron, A. (eds.) RSFDGrC 2003. LNCS (LNAI), vol. 2639, pp. 274–278. Springer, Heidelberg (2003)
12. Ziarko, W., Wong, S.K.M.: On optimal decision rules in decision table. Bulletin of Polish Academy of Sciences 33(11-12), 693–696 (1985)
13. Wang, P.-C.: Highly scalable rough set reducts generation. Journal of Information Science and Engineering 23(4), 1281–1298 (2007)
14. Herbert, J.P., Yao, J.T.: Game-theoretic rough sets. Fundamenta Informaticae 108(3-4), 267–286 (2011)
15. Yao, J.T., Herbert, J.P.: A game-theoretic perspective on rough set analysis. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition) 20(3), 291–298 (2008)
16. Herbert, J.P., Yao, J.: Analysis of Data-Driven Parameters in Game-Theoretic Rough Sets. In: Yao, J., Ramanna, S., Wang, G., Suraj, Z. (eds.) RSKT 2011. LNCS, vol. 6954, pp. 447–456. Springer, Heidelberg (2011)
17. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944)
18. Herbert, J.P., Yao, J.T.: A Game-Theoretic Approach to Competitive Learning in Self-Organizing Maps. In: Wang, L., Chen, K., S. Ong, Y. (eds.) ICNC 2005. LNCS, vol. 3610, pp. 129–138. Springer, Heidelberg (2005)
19. Azam, N., Yao, J.T.: Incorporating Game Theory in Feature Selection for Text Categorization. In: Kuznetsov, S.O., Ślęzak, D., Hepting, D.H., Mirkin, B.G. (eds.) RSFDGrC 2011. LNCS, vol. 6743, pp. 215–222. Springer, Heidelberg (2011)
20. Kol, G., Naor, M.: Cryptography and Game Theory: Designing Protocols for Exchanging Information. In: Canetti, R. (ed.) TCC 2008. LNCS, vol. 4948, pp. 320–339. Springer, Heidelberg (2008)
21. Saad, W., Han, Z., Debbah, M., Hjørungnes, A., Basar, T.: Coalitional game theory for communication networks. IEEE Signal Processing Magazine 26(5)
22. Yao, Y.: Decision-Theoretic Rough Set Models. In: Yao, J., Lingras, P., Wu, W.-Z., Szczuka, M.S., Cercone, N.J., Ślęzak, D. (eds.) RSKT 2007. LNCS (LNAI), vol. 4481, pp. 1–12. Springer, Heidelberg (2007)
23. Yao, Y.Y., Zhao, Y., Wang, J.: On reduct construction algorithms. In: Wang, G.-Y., Peters, J.F., Skowron, A., Yao, Y. (eds.) RSKT 2006. LNCS (LNAI), vol. 4062, pp. 297–304. Springer, Heidelberg (2006)
24. Swiniarski, R.W., Skowron, A.: Rough set methods in feature selection and recognition. Pattern Recognition Letters 24(6), 833–849 (2003)

Context-Based Decision Support for Energy Efficiency and Emissions Trading in Industry: The LifeSaver Approach

Rui Neves-Silva, Maria Marques, Ana Rita Campos, Christian Wolff, and Boris Sucic

Abstract. This paper reports the work under development in the scope of LifeSaver project aiming at supporting manufacturing companies in optimising the energy performance of their operations. This support will be in the form of a set of ICT building blocks that combine context awareness, ambient intelligence monitoring and standard energy consumption data measurement. LifeSaver will provide: (i) information about the energy consumption to be processed in enterprise management systems, for the purpose of achieving significant energy savings; (ii) a decision support system for optimisation of energy performance of operations; and (iii) predicted cumulative data on the CO2 emissions, as input for the decision support services to enable emission trading across industries and among companies. Three demonstrators of the platform's application in real industrial environment, and their usage for new business models, will be provided.

1 Introduction

This paper reports the work being developed under the LifeSaver project [1]; a European project with 3 years duration that started in December 2011. LifeSaver

Rui Neves-Silva
DEE-FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
e-mail: rns@fct.unl.pt

Maria Marques · Ana Rita Campos
UNINOVA, FCT Campus, 2829-516 Caparica, Portugal

Christian Wolff
ATB- Bremen, Wiener Strasse 1, 28359 Bremen, Germany

Boris Sucic
Jozef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia

J. Watada et al. [Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 185–194.
springerlink.com

project aims at supporting manufacturing companies in optimising the energy performance of their operations.

In the last decade, increased attention has been given to energy efficiency and greenhouse gas emissions. According to the International Energy Agency, global energy demand will grow 55% by 2030. In the period up to 2030, the energy supply infrastructure worldwide will require a total investment of USD 26 trillion, with about half of that in developing countries. If the world does not manage to green these investments by directing them into climate-friendly technologies, emissions will go up by 50% by 2050, instead of down by 50%, as science requires [2]. The European Commission has prioritized climate change and energy in its new Europe 2020 strategy, establishing ambitious goals: 20% reduction of greenhouse gas emissions, meeting 20% of energy needs from renewable sources, and reducing energy consumption by 20% through increased energy efficiency.

The European Emission Trading Scheme (ETS), introduced by EU ETS Directive (2003/87/EC) is in the second trading period (2008 – 2012)[3]. Industrial installations participating in the ETS are obliged to monitor and annually report their emissions and surrender enough allowances to cover all their emissions at the end of each year. The third trading period (2013 - 2020) will bring significant changes for industrial installation as number of free allocated allowances will be significantly decreased and half of the total number of allowances is expected to be auctioned. This will bring new dynamics in the emission trading activities to mitigate economic risks in industrial companies. The use of intelligent tools for efficient management of energy and emission flows in the company, will enable effective trading processes on different types of actions (yearly, monthly, daily) and use of other flexible mechanisms (Joint implementation and Clean Development Mechanism) to fulfill efficiently their emission obligations.

However, emissions trading is rather new process/concept in everyday business for companies that are required by law to procure coupons for their CO2 emissions. As such it is not their primary business process and so not their main focus when managing their primary business process. For companies that produce high amounts of CO2 emissions this could mean that money will be spent on buying emission coupons which could be avoided with better planning/optimization. Since companies that are large CO2 emissions producers most probably already have some sort of EMS (Energy Management System) for managing their energy consumption an improved system is needed to analyze and advise how to manage emissions based on the EMS data and potential energy efficiency projects reducing CO2 emissions.

LifeSaver project is not intended to develop an emissions trading system per se. It is intended as a Decision Support System taking into account the specific energy usage pattern of a company and monitoring and control of energy and other parameters influencing emissions. Using historical values of CO2 emissions of a company, LifeSaver includes functionalities to predict future emissions which will then be the basis for the recommendations on how to trade with emissions coupons.

In particular, manufacturing processes are increasingly complex and flexible, leading to highly dynamic energy usage patterns and difficulties in identification

of correlation between CO2 emissions and specific processes. To cope with such a high dynamic of energy patterns and to enable an efficient energy management in manufacturing processes, LifeSaver complements currently measured energy consumption data with diverse information from ambient intelligent systems (e.g. interactions between human operators and machines or processes) and process-related measurements (e.g. already available for control or monitoring purposes). The main objective is to enrich energy consumption data with information about the context in which the energy consumption actually occurred. This enables LifeSaver's building blocks to access and process more complete information to target energy efficiency optimization and CO2 emission reduction.

There are several methods developed to support human actors in making decisions, in several application fields. When considering a scenario of saving energy, the decision process has to consider a diverse range of criteria, covering economics, environment, operation, logistics, etc. To address such processes, the users need a method that addresses multi-criteria decision analysis or multi-criteria decision making. This scientific topic supports human actors in evaluating situations, highlighting conflicts and deriving a compromise to make a decision. The current top multi-criteria decision methods are: aggregated indices randomisation method, analytic hierarchy process, analytic network process, ELECTRE, goal programming, data envelope analysis, dominance-based rough set approach, grey relational approach, multi-attribute utility theory, multi-attribute value theory, PROMETHEE, MACBETH, value analysis, weighted sum and weighted product model. Some of these methods intend to select the best choice (e.g. multi-attribute utility theory) and others provide an ordered set of alternatives (e.g. AHP, ANP) [4].

Some of these methods have been realised into decision support systems, i.e. interactive computer-based information systems designed to help human decision makers. These systems process data and models in order to identify, structure, and solve semi-structured or unstructured problems [5],[6]. More recently, there was a convergence among the research fields of decision support, artificial intelligence and knowledge management. From this, a number of information systems emerged to generate knowledge for decision-making support. Collectively, these systems can be called intelligent decision support systems [7]. These systems integrate the functions of decision support systems and knowledge based systems to assist decision makers in building analytical models, offer advice on specific problem tasks, assist decision makers in performing decision analysis, and explain conclusions and recommendations [8],[9],[10],[11].

There are some methods and systems developed in the application area of energy, but they mainly focus on energy planning and policy analysis. These systems address issues such as selecting the best renewable energy source for a region, planning energy distribution, selecting location for wind generators or other technologies. These systems are not suitable for daily operation in industrial companies. LifeSaver helps human actors make daily operational and business decisions, while considering an energy savings perspective. Additionally, LifeSaver considers the several sources of uncertainty in the criteria and indicators, which is something not sufficiently addressed in existing approaches.

2 Concept and Objectives

The project aims at supporting manufacturing companies in optimizing energy performance of their operations through a set of ICT building blocks that combine context awareness, ambient intelligence monitoring and standard energy consumption data measurement. LifeSaver provides:

- comprehensive information about the energy consumption to be processed in EMS, for the purpose of achieving significant energy savings,
- a knowledge-based decision support system for optimization of energy performance of operations, and
- appropriate (almost) online and predicted cumulative data on the CO2 emissions, as input for the decision support services to enable emission trading.

The set of building blocks allows for identification of energy profiles and energy consumption patterns and their interrelations to support energy efficiency optimization and CO2 emissions reduction and trading, using the above mentioned approach of combining knowledge management and ambient intelligence technology for innovative data/information/knowledge processing.

The monitored energy consumption enriched with context from ambient intelligence data, can be the basis for the identification of energy profiles and energy consumption and emissions patterns. These profiles and patterns are the basis to support decision-making in different time-frame situations (see Figure 1.):

a) short-term routine decisions related to daily use of equipment, aiming at maximising energy savings;

b) medium-term operational decisions concerning the re-configuration of processes in manufacturing companies, aiming at improving energy-efficiency, and

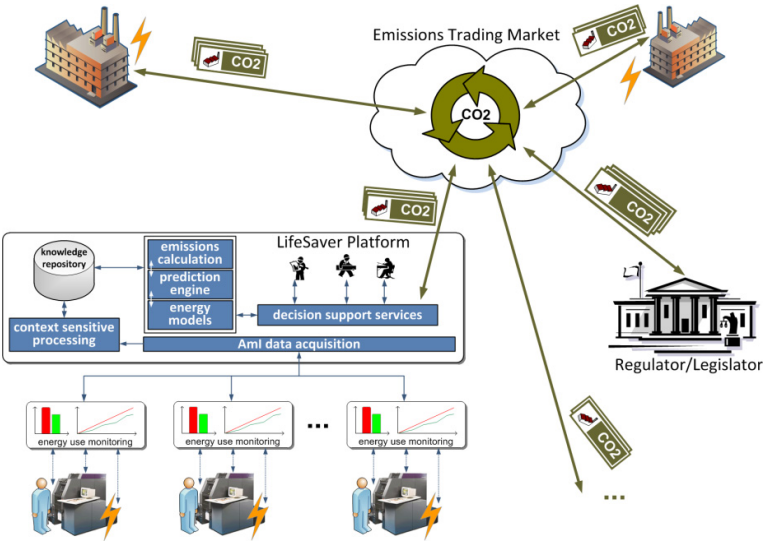


Fig. 1 The proposed concept.

c) long-term decisions aiming at managing cumulative emissions and supporting emissions trading among companies.

LifeSaver's building blocks will comprise blocks with the following functionality :

- **AmI data acquisition:** acquires the monitored energy consumption measuring data and information about the manufacturing processes in order to provide context sensitive data to monitor energy use and calculate CO2 emissions.
- **Context sensitive processing:** enriches the monitored data with information already stored in the knowledge repository or also measured, to provide insight about how the energy was used, i.e. connect monitored data with context under which the energy was used, and enables the identification of warnings related to excessive energy consumption.
- **Energy models:** elaborate the energy profile of a company or a section of the company, combining individual data on machines or processes.
- **Prediction engine:** uses the measurements acquired and the energy models to forecast the energy consumption of the company for a defined upcoming period (e.g. predict the energy consumption of the complete company for a whole year based on individual machine measurements of some weeks).
- **Emissions calculation:** uses the energy consumption measured and elaborated energy models to calculate the CO2 emissions generated by the company and extrapolate the cumulative emissions for a defined period of time.
- **Decision support services:** provide systematic mechanism for short and medium-term decisions on how to save energy in operations and long-term decisions on how to reduce emissions in a company and also on how to trade emissions among companies.

This platform should be open for integration with other enterprise management systems.

3 Decision-Making Process

The decision support services are being developed in line with the approach taken to the prediction engine. The prediction models serve the decision services by providing means to build and test possible future scenarios. These scenarios, i.e. decision alternatives, will be scored in energy consumption emissions criteria according with the outputs of the prediction engine. Additionally, other criteria (e.g. cost impact) or hard constraints (e.g. safety regulations) may apply. Being a decision support system, the human actor(s) play a central role in the decision process, even by complementing the set of scenarios with additional empiric information or additional scenarios, and validating the final decision to implement. LifeSaver provides decision support functionalities to automatically build up scenarios, order alternatives on cost-benefit, perform sensitive analysis (e.g. safety impact over the small variation of a decision parameter), decision profiling, logging and reporting. Here, as well, the context under which the decision process is occurring will push the access to higher levels of knowledge, not usually available.

The key challenge to the practical success of the proposed approach is to make these services as seamless as possible, supporting the user interactions without requiring actors to explicitly document all their actions [12],[13].

To illustrate the basis of the decision process, consider an industrial company with several similar units producing only two different types of products: *type 1* and *type 2*. Additionally, consider that because of business operations/constraints the units have been used in different production rates, from nominal to high. Figure 2 depicts a map of cases obtained from AmI data acquisition from different time periods and different units. In this simple case, only two variables define the context: *production rate* and *type of product* being produced. The different marks in the picture represent the consumption intensity per product in different machines.

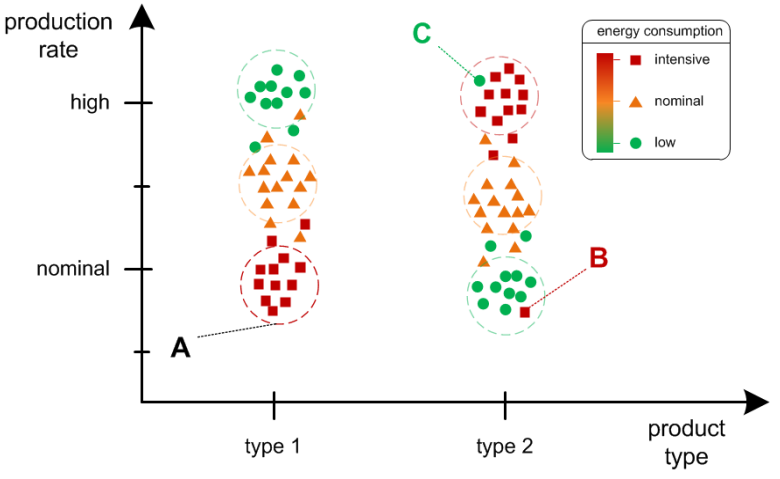


Fig. 2 Example on map of the energy consumption according with production context.

The proposed approach can recognize, from the analysis of this data, three different situations that are to be processed by the LifeSaver system:

A. The direct correlation between context variables and energy intensity. For instance, we can identify a cluster of records around product *type 1*, under a nominal production ratio, with an intensive use of energy (marked A in Figure 2).

   The processing of this knowledge will continuously build adaptive energy models describing the general possible operating points in the context envelop, regarding energy use. The energy models are defined by a collection of cases reporting expected consumption patterns.

B. From current on-line acquired AmI data, the system identifies a "negative" outlier record (marked B in Figure 2), since for the current context of the machine, it should exhibit a low consumption pattern.

C. In opposition to the previous case, the system identifies a current "positive" outlier record (marked C in Figure 2) where an intensive consumption was expected from the energy models being identified.

Both situations B. and C. are to be identified using Case-Based Reasoning. The context is continuously identified and compared with the expected consumption patterns previously obtained (situation A. above). When the these patterns deviate from the "normal" within that context, the system searches for similar patterns (past cases), cluster the cases and provides suggestions to the user (e.g. maintenance teams) on the best action to correct (situation B.) or re-use in similar machines (situation C.).

In parallel with the continuous AmI data acquisition and context sensitive processing, the decision support services are available to the end-user, making use of the knowledge to provide:

1. Support for short-term problem identification and solving from typical cases described in item B above. The service identifies the equipment and supports the user to understand if the abnormal situation, with increased consumption, results from an equipment mal-function or from an additional context, not fully observable using the current set of AmI sensors and context processing items.
2. Support for medium-term reconfiguration of system (e.g. controllers) to explore energy-efficient opportunities identified either directly from records as the one described in item C above, either by proactive field trials against the energy models referred in item A above.
3. Support for long-term (e.g. 1 year) prediction of future energy consumption and corresponding calculated CO2 emissions, taking into consideration the future business objectives/constraints in cooperation with the company's EMS, to enable production planning and emissions trading. Here, the decision support approach is cost oriented.

The complexity of such system, as described above, will arise from the many variables that will provide contextual information – some relevant, some not – including human interaction during operations. Moreover, the decision process will eventually derive in actions (e.g. reconfiguration of the equipment) and this will also be part of the context of past records stored in the repository.

From these use case scenarios, the following technical challenges are identified:

- How to measure and process the context of industrial operations to support elaboration of adequate energy models?
- How to adequately model the energy patterns to support energy consumption monitoring and prediction?
- How to accurately derive the emissions in an extended horizon to support companies' decisions in a trading market?
- How to combine short-term and long-term decision making to support business objectives while reducing energy consumption, emissions and overall costs?

LifeSaver will develop the necessary methods to answer these research questions that will provide the means to specify and develop the LifeSaver platform.

4 Main Results

To achieve the stated objectives the project will deliver the following:

- the LifeSaver platform, comprehending a collection of modules and services to support decisions that enable the optimization of industrial operations, improving their energy performance;
- the LifeSaver methodology includes a set of methods aimed at supporting RTDs and technology providers in developing work in the following topics:
  - AmI data acquisition and context-aware processing;
  - Energy models and prediction;
  - Emissions calculation and trading;
  - Decision support for energy savings and emissions trading;
- application of the platform within three demonstrators:
  - one focusing on short-term daily decisions;
  - another focusing on medium-term reconfiguration of equipment; and
  - the third one focusing on long-term emissions management and trading.

The LifeSaver platform will be built to be easily applicable with existing repositories and other legacy systems in industrial companies.

5 Applications

The LifeSaver consortium comprehends eight partners. Two of these organizations are technology providers, which will support the development and validation of results, particularly regarding the IT development and integration with control systems. The consortium also includes two industrial end-user companies and one municipal energy agency to provide testing environments for the project's results.

The tools will be applied and validated in real infrastructures in the scope of the following three scenarios:

- The first includes two German companies, where one supplies control systems to the operations of the second. The objective of this business case scenario is to maximize the potential of energy savings in the second industrial company, and explore how the control systems provider can influence those savings from the project development stage.
- The second scenario is centered around a Slovenian enterprise with a very significant energy consumption profile needing to implement a systematic approach to save energy and also calculate generated emissions.
- Finally, a Portuguese municipal energy agency has the objective of contributing to the sustainable development of the people and companies of its region. This business case will test by simulation an emissions trading market with several companies in the region.

The project aims at supporting companies in reducing their energy consumption in 10% to 15% (depending on the initial conditions and specific processes) and, consequently, approaching and surpassing the benchmarked emissions specified in the EC decision for free allocation of emissions allowances (European Commission 2010). One of the key results by which the process results will be validated is the support to the companies in trading emissions. The industrial end-users in the consortium are already well aware of their total energy consumption and emissions relatively to the benchmarks for their sectors. The objective will be to break down the energy model of each company in sectors enabling the verification of more focused results, i.e. related with specific products, processes or even machines. The AmI-based monitoring that industrial end-users will apply will allow developing these high resolution energy profiles. In the beginning of the project, the main verification indicators are the total energy consumption and corresponding emissions. Nevertheless, throughout the project, with the additional amount of information, more detailed indicators per company activity will become available and specific potentials for reduction of energy consumption will be precisely defined.

6 Conclusions and Future Work

The LifeSaver approach is based on the monitoring of the energy performance of operations in manufacturing companies. The collected contextualized knowledge will be used to support companies in reconfiguring its operations to optimize the participation in the emissions trading market.

The project aims at supporting companies in reducing their energy consumption in 10% to 15% (depending on the initial conditions and specific processes) and, consequently, approaching and surpassing the benchmarked emissions specified in the draft EC decision for free allocation of emissions allowances [14].

Acknowledgment. Authors express their acknowledgement to the consortium of the project LifeSaver, *Context sensitive monitoring of energy consumption to support energy savings and emissions trading in industry* [http://www.lifesaver-fp7.eu]. LifeSaver is funded under the Seventh Research Framework Program of the European Union (contract FP7-ICT-7-287652).

References

1. LifeSaver, Consortium. LifeSaver Project, `http://www.lifesaver-fp7.eu`
2. United Nations Framework Convention on Climate Change. Fact Sheet: The need for strong global action on climate change. United Nations Framework Convention on Climate Change (February 2011), `http://unfccc.int/2860.php` (accessed 2011)
3. European Commission. Legislation on Emissions Trading System, `http://ec.europa.eu/clima/policies/ets/documentation_en.htm#Main_legislation`

4. Georgilakis, P.S.: State-of-the-Art of decision support systems for the choice of renewable energy sources for energy supply in isolated regions. International Journal of Distributed Energy Resources 2(2), 129–150 (2006)
5. Sprague, R.H.J.: A framework for the development of decision support systems. MIS Quaterly 4(4), 1–26 (1980)
6. Zolghadri, M., Lecompte, T., Bourrieres, J.P.: Design of decision support systems for extended enterprise. Journal of Studies in Informatics and Control with Emphasis on Useful Applications of Advanced Technology 11(1)
7. Hans, W.G., Peter, W.: Intelligent decision support systems. Decision Support Systems 8(4), 317–332 (1992)
8. Hunsaker, P.L., Hunsaker, J.S.: Decision styles - in theory, in practice. Organizational Dynamics 10, 23–26 (1981)
9. Tansley, D.S.W., Hayball, C.C.: KBS Analyst and Design, a KADS Developer's Handbook. Prentice-Hall (1993)
10. Sensiper, S.: American Management Systems: The Knowledge Centers. Case 9-697-068. Harvard Business School (1998)
11. Silverman, B.G.: Unifying expert systems and the decision sciences. Operations Research 42(3), 393–413 (1994)
12. Mazharsolook, E., Scholze, S., Neves-Silva, R., Ning, K.: Enhancing networked enterprise management of knowledge and social interactions. Journal of Computing in Systems and Engineering, 176–184 (2009)
13. Campos, A.R.: Intelligent Decision Support Systems for Collaboration in Industrial Plants. PhD Thesis (2010)
14. European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Addressing the challenge of energy efficiency through information and communication technologies (2008)

Entropy Model of a Fuzzy Random Portfolio Selection Problem

Takashi Hasuike and Hideki Katagiri

Abstract. This paper considers an entropy model of portfolio selection problem with fuzzy random variables to future returns. Since standard mean-variance portfolio models suffer from some shortcomings, the entropy is introduced as a risk measure instead of variances to overcome the shortcomings. Furthermore, introducing the sum of entropy to each portfolio as well as the entropy of fuzzy random variables, the previous entropy-based fuzzy random portfolio selection problem is extended, the exact optimal portfolio is explicitly obtained using nonlinear programming such as Karush-Kuhn-Tucker condition.

1 Introduction

Portfolio selection problem is regarded as the decision of optimal asset allocation to securities among various securities, and it is one of the most important research themes in financial and investment research fields. Until now, many researches concerning portfolio selection problems have been studied. Regarding the research history on mathematical approaches, Markowitz [14] proposed the mean-variance model in which each variance of returns is dealt with a risk factor. Since then, it has been the centre of research activities in the real investment field, and numerous researchers have contributed to the development of modern portfolio theory. Many researchers have proposed mathematical models for portfolio selection problems that have extended the Markowitz model; mean-absolute-deviation model

Takashi Hasuike
Graduate School of Information Science and Technology, Osaka University, Japan,
2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan
e-mail: thasuike@ist.osaka-u.ac.jp

Hideki Katagiri
Graduate School of Engineering, Hiroshima University, Japan, 1-4-1 Kagamiyama,
Higashi-Hiroshima, Hiroshima, 739-8527 Japan
e-mail: katagiri-h@hiroshima-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 195–203.
springerlink.com

(Konno et al. [10, 11]), semi-variance model (Bawa and Lindenberg [2]), safety-first model (Elton and Gruber [5]), Value-at-Risk (VaR) and conditional Value-at-Risk (cVaR) model (Rockafellar and Uryasev [19]) etc..

In these existing portfolio models, they assumed that future return occurs according to a random distribution derived from statistical analysis of market data. On the other hand, investors receive not only effective but also ineffective information from the real market in which ambiguous factors usually exist. Furthermore, investors often have their subjective predictions for future markets, and these subjectivities are not derived from the statistical analysis of historical data, but from their long-term experiences. Even if investors hold a lot of information from current financial markets, it is difficult that the present or future random distribution of each asset is strictly set. Consequently, we need to consider not only random conditions but also ambiguous and subjective conditions for portfolio selection problems.

As recent studies in mathematical programming, some researchers have proposed various types of portfolio models under randomness and fuzziness. These problems with probabilities and possibilities are generally called stochastic programming problems and fuzzy programming problems, respectively, and there are some basic studies using a stochastic programming approach, goal programming approach, etc.., and fuzzy programming approach to treat ambiguous factors as fuzzy sets (Inuiguchi and Ramik [7], Leon, et al. [13], Tanaka and Guo [22], Tanaka, et al. [23], Vercher et al. [24], Watada [25]). Furthermore, some researchers have proposed mathematical programming problems with both randomness and fuzziness as fuzzy random variables (for instance, Katagiri et al. [8, 9]). In the studies [8, 9], fuzzy random variables were related with the ambiguity of the realization of a random variable and dealt with a fuzzy number that the center value occurs according to a random variable. The uncertainty leads to different people having different feelings on observing the same event, and so fuzzy random portfolio selection problems may be more useful than previous portfolio models with only random or fuzzy variable.

Almost above-mentioned portfolio models are based on or extend to the mean-variance model. The main concept of mean-variance model is to regard each variance as the risk measure. However, in general, using variance to measure the portfolio risk may suffer from some shortcomings. For instance, in the mean-variance model, some strict assumptions have to be imposed on security return, i.e., it requires that securities' return rates obey normal distribution or the investors have binomial utility function, which can hardly be satisfied in real security market. Furthermore, calculating the variance covariance matrix for a portfolio with large number of securities is a huge computation task. As a risk measure to overcome these shortcomings from variance-based risk modeling approaches, some researchers introduce an entropy into portfolio selection problems (Philipatos and Wilson [16], Popkov [17], Simonelli [20], Smimou et al. [21], Yang and Qiu [27]). These previous entropy-based models can be only employed when probability distribution is assumed, and so a new entropy measure is required when one studies hybridized uncertainty of both fuzziness and randomness in modern complicated financial market. In most recently, Xu et al. [26] have proposed a hybrid entropy as a measure of

security risk which could describes both random and fuzzy uncertainty, and formulated some hybrid entropy-based fuzzy random portfolio selection problems. However, these models are linear programming problems, and so the optimal asset allocation tends to concentrate a few specific securities. In a numerical example of Xu et al. [26], they considered ten securities as the investment, but the optimal asset allocation concentrated to only two securities. In Xu et al. [26], they considered only the entropy to fuzzy random variables of future returns. In the investment, the diversification is one of the most important factors to avoid the risk. Therefore, in this paper, by introducing the entropy to the sum of each portfolio as well as fuzzy random variables of future returns, we extend the previous entropy-based fuzzy random portfolio selection problem.

Since our model is a bi-objective programming problem with both entropies to the sum of each portfolio and fuzzy random variables of future returns, in this paper, we transform main problems into the deterministic equivalent problems introducing an aggregated function with an weight to each object decided by the investor, and explicitly obtain the exact optimal portfolio using nonlinear programming approach such as Karush-Kuhn-Tucker (KKT) condition.

This paper is organized as follows. In Section 2, we introduce the basic mathematical knowledge of fuzzy random variables, λ mean ranking method, and hybrid entropy to fuzzy random variables. In Section 3, using the hybrid entropy and standard entropy to the sum of each portfolio, we propose a new entropy-based fuzzy random portfolio selection problem extending the study of Xu et al. (2011), and develop the solution algorithm to obtain an exact optimal portfolio using the KKT condition. Finally, in Section 4, we conclude this paper and discuss future researches.

2 Preliminaries

2.1 Fuzzy Random Variable

Consider a fuzzy set theory before introducing the definition of a fuzzy random variable. The fuzzy set theory was proposed by Zadeh [28] as a means of representing and manipulating data that was not precise, but rather fuzzy, and it was specifically designed to mathematically represent uncertainty and vagueness. Therefore, it allows decision making with estimated values under incomplete or uncertain information (Carlsson and Fuller [4]). The mathematical definition of fuzzy sets is given as follows.

Definition 1. Let $\hat{X}$ be a nonempty set. A fuzzy set Φ in $\hat{X}$ is characterized by its membership function $\mu_\Phi : \hat{X} \to [0,1]$ and μ_Φ is interpreted as the degree of membership of element x in fuzzy set Φ for each $x \in \hat{X}$.

With respect to the degree to which the statement gx is in Φh is true, Definition 1 means that the value zero is used to represent complete non-membership, the value one is used to represent complete membership, and values in between are used to represent intermediate degrees of membership.

Using this fuzzy sets theory as well as the probability theory, a fuzzy random variable was first defined by Kwakernaak [12], and Puri-Ralescu [18] established the mathematical basis. In this paper, consider the case where the realization of random variable is a fuzzy set. Accordingly, fuzzy random variables are defined as follows:

Definition 2. Let (Ω, B, P) be a probability space, $F(\Re)$ be the set of fuzzy numbers with compact supports, and $\hat{X}$ be a measurable mapping $\Omega \to F(\Re)$. Then, $\hat{X}$ is a fuzzy random variable if and only if given $\omega \in \Omega, \hat{X}_h(\omega)$ is a random interval for any $h \in [0,1]$ where $\hat{X}_h(\omega)$ is a h-level set of the fuzzy set $\hat{X}(\omega)$.

The above definition of fuzzy random variables corresponds to a special case of those given by Kwakernaak and Puri-Ralesu. The definitions of them are equivalent for the above case because a fuzzy number is a convex fuzzy set. Though it is a simple definition, it would be useful for various applications.

2.2 *Expected Value of Fuzzy Random Variable with λ Mean Ranking Method*

In the case we compare the fuzzy random variables, it is standard and simple that we compare the expected values of fuzzy random variables. Some researchers have proposed the mathematical formulation of expected values of fuzzy random variables (for instance, Gil and Lopez-Diaz [6], Puri and Ralescu [18]). To simplify, if we calculate an expected value of fuzzy random variable in probability theory, the expected value is represented as a fuzzy number. Therefore, we have to induce order relations to compare and rank fuzzy numbers. In this paper, we use the λ mean ranking method proposed by Campos and Gonzalez [3] to deal with the fuzzy decision-making problem. This method allows us to introduce the decision maker (DM)'s subjective opinion, i.e. the pessimistic-optimistic parameter λ in the decision making process.

Definition 3. Let $\tilde{M}$ be a fuzzy number, its α-cut set is $\tilde{M}_\alpha = [M_\alpha^-, M_\alpha^+]$. The fuzzy number $\tilde{M}$'s λ mean is defined as

$$E^\lambda\left(\tilde{M}\right) = \int_0^1 \left(\lambda M_\alpha^+ + (1-\lambda) M_\alpha^-\right) d\alpha \tag{1}$$

where the parameter $\lambda \in [0,1]$ reflects the DM's subjectivity to the optimistic-pessimistic degree.

The parameter $\lambda = 1$ expresses that DM is totally optimistic about all of the maximal value (delegate the maximal possible return). On the contrary, the parameter $\lambda = 0$ expresses that DM is totally pessimistic (delegate the minimal possible return). If the parameter $\lambda = 0.5$, it expresses that the DM's attitude is relative neutral. In general, if the DM is relative optimistic, then the parameter could be a little larger number, and if the DM is relative pessimistic, then the parameter could be a little smaller number. Thus, through changing the value of parameter $\lambda \in [0,1]$, the DM's optimistic-pessimistic degree can be reflected.

2.3 Hybrid Entropy for Fuzzy Random Variable

As discussions in Introduction section, complicated financial instruments can exhibits hybridized uncertainty of both randomness and fuzziness. In the study of Xu et al. [26], they introduced a extend entropy called Hybrid entropy to measure financial risk caused by both randomness and fuzziness, simultaneously. The definition of hybrid entropy is as follows:

Definition 4. Let $D(X)$ be the collectivity of sets on discrete universe of discourse $X = \{x_i | i = 1,2,...,n\}$, $A \in D(X)$, denoted as $A = \sum_{i=1}^{n} \mu_i / x_i$, $\forall x_i \in X$, wherein $\mu(x_i)$ is the membership of x_i subjects to A, and the range is $[0,1]$. Let $M = (\mu_1, \mu_2, ..., \mu_n)$, called the possibility distribution of fuzzy set A, and the probability distribution on the universe of discourse X is $P = (p_1, p_2, ..., p_n)$, where p_i is the frequency of x_i. We define the hybrid set as $\bar{A}(P,M) = \sum_{i=1}^{n} p_i \mu_i / x_i, \forall x_i \in X$, recorded as $\bar{A}(P,M) = \bar{A}$. Then, we call

$$F(\bar{A}) = -\sum_{j=1}^{n} \{p_j \mu_j \log p_j \mu_j + p_j (1-\mu_j) \log p_j (1-\mu_j)\} \tag{2}$$

the hybrid entropy of the hybrid set $\bar{A}$.

This hybrid entropy is also called Total entropy, some researchers have proposed the same or equivalent transformed mathematical formulation and considered the application (for instance, Al-sharhan et al. [1], Nieradka and Butkiewicz [15]). Using this hybrid entropy, Xu et al. [26] proposed a more reasonable and extensive portfolio selection model satisfying the different investors' attitude and experts' opinions. However, their models are formulated as linear programming problems, and so the optimal asset allocation tends to concentrate a few specific securities, because they considered only the entropy to fuzzy random variables of future returns. Therefore, to achieve the diversification, we extend the previous entropy-based fuzzy random portfolio selection problem by introducing the entropy to the sum of each portfolio as well as fuzzy random variables of future returns.

3 Extended Entropy-Based Fuzzy Random Portfolio Selection Problem

3.1 Formulation of Our Proposed Model

Let us consider fuzzy random future returns $\tilde{\bar{r}}_j$, which is assumed that the realization for the random variable is a fuzzy number $\tilde{r}_j(\omega)$ characterized by the following triangle membership function $\tilde{r}_j(\omega) = (r_j(\omega), \alpha_j, \beta_j)$:

$$\mu_{\tilde{\bar{r}}_j(\omega)}(\xi) = \begin{cases} \dfrac{\xi - (r_j(\omega) - \alpha_j)}{\alpha_j} & (r_j(\omega) - \alpha_j < \xi \le r_j(\omega)) \\ \dfrac{(r_j(\omega) + \beta_j) - \xi}{\beta_j} & (r_j(\omega) < \xi \le r_j(\omega) + \beta_j) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $r_j(\omega)$ is the realization of a random variable $\bar{r}_j$ whose mean is m_j and variance is σ_j^2. Then, parameters α_j and β_j are positive values representing the left and right spreads of fuzzy numbers, respectively. The triangle membership function is a special case of general L-R fuzzy number, and it is used as a general membership function in many practical situations.

Furthermore, we assume the following two situations:

(a) The hybrid entropy (2) to fuzzy random variables denotes the average information value of the security. If $p(r_j) = 0$ or $\mu(r_j) = 0$, the limit value of $F(\tilde{\bar{r}}_j)$ is 0. In the case of a positive value, a larger value of $F(\tilde{\bar{r}}_j)$ means greater degree of uncertainty of obtaining the return, i.e., higher risk. On the contrary, in the case of negative value, a smaller value of $F(\tilde{\bar{r}}_j)$ means greater degree of uncertainty of obtaining the return, i.e., higher risk. Thus, it is obvious that a larger value of $F(\tilde{\bar{r}}_j)$ indicates a higher risk, and so the first object of our proposed model is minimizing $\sum_{j=1}^{n} x_j F(\tilde{\bar{r}}_j)$ with each hybrid entropy of fuzzy random returns $F(\tilde{\bar{r}}_j)$, where x_j is the portfolio rate of security j.

(b) Since the diversification is also one of the most important factors to avoid the risk, the second object is maximizing the standard entropy to the sum of each portfolio rate formulated as $-\sum_{j=1}^{n} x_j \log x_j$.

From these two assumptions, we formulate a new entropy-based fuzzy random portfolio selection problem as follows:

$$\begin{aligned} &\text{Minimizing} \quad \sum_{j=1}^{n} x_j F(\tilde{\bar{r}}_j) \\ &\text{Maximizing} \quad -\sum_{j=1}^{n} x_j \log x_j \\ &\text{subejct to} \quad E\left(\sum_{j=1}^{n} \tilde{\bar{r}}_j x_j\right) \geq r_G, \\ &\qquad \sum_{j=1}^{n} x_j = 1, x_j \geq 0 \end{aligned} \tag{4}$$

where r_G is the target value of total expected future return. We denote that parameter vector of optimistic-pessimistic degree $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$ is decided by the investor. Then, from Definition 3 and triangle fuzzy number (3), the λ-mean of $E\left(\sum_{j=1}^{n} \tilde{\bar{r}}_j x_j\right)$ is calculated as follows:

$$\begin{aligned} E^{\lambda}\left(\sum_{j=1}^{n} \tilde{\bar{r}}_j x_j\right) &= \sum_{j=1}^{n} \left(E(\bar{r}_j) + \lambda_j \beta_j \int_0^1 (1-\alpha)\, d\alpha - (1-\lambda_j)\alpha_j \int_0^1 (1-\alpha)\, d\alpha\right) x_j \\ &= \sum_{j=1}^{n} \left(m_j + \frac{1}{2}(\lambda_j \beta_j - (1-\lambda_j)\alpha_j)\right) x_j = \sum_{j=1}^{n} \hat{m}_j(\alpha_j, \beta_j, \lambda_j) x_j \end{aligned} \tag{5}$$

Furthermore, our proposed model is a bi-objective programming problem, and so it is hard to solve this problem directly in mathematical programming. In this paper, we introduce an aggregated function with a weight w to each object decided by the investor satisfying $0 \leq w \leq 1$, and reformulate our proposed model as the following single objective programming problem:

$$\begin{aligned} &\text{Minimizing} \quad (1-w)\sum_{j=1}^{n} x_j F(\tilde{\bar{r}}_j) + w\sum_{j=1}^{n} x_j \log x_j \\ &\text{subejct to} \quad \sum_{j=1}^{n} \hat{m}_j(\alpha_j,\beta_j,\lambda_j)x_j \geq r_G, \\ &\qquad \sum_{j=1}^{n} x_j = 1, x_j \geq 0 \end{aligned} \tag{6}$$

3.2 Development of Exact Solution Algorithm

In problem (6), since $\sum_{j=1}^{n} x_j \log x_j$ is a convex on each x_j and the others are linear functions, problem (6) is a convex programming problem. Therefore, using nonlinear programming approaches, we obtain a global optimal solution. To simplify, we assume the case $\sum_{j=1}^{n} \hat{m}_j(\alpha_j,\beta_j,\lambda_j)x_j = r_G$. We apply the method of Lagrange multipliers and Karush-Kuhn-Tucker (KKT) condition. The Lagrange function for problem (6) is as follows:

$$\begin{aligned} L = &(1-w)\sum_{j=1}^{n} x_j F(\tilde{\bar{r}}_j) + w\sum_{j=1}^{n} x_j \log x_j \\ &+ \gamma\left(r_G - \sum_{j=1}^{n} \hat{m}_j(\alpha_j,\beta_j,\lambda_j)x\right) + \eta\left(1 - \sum_{j=1}^{n} x_j\right) - \sum_{j=1}^{n} u_j x_j \end{aligned} \tag{7}$$

where γ, η, and u are the Lagrange multiplier. Then, KKT condition of this Lagrange function is obtained as follows:

$$\begin{cases} \dfrac{\partial L}{\partial x_j} = (1-w)F(\tilde{\bar{r}}_j) + w(\log x_j + 1) \\ \qquad\qquad - \gamma\hat{m}_j(\alpha_j,\beta_j,\lambda_j) - \eta = 0, (j = 1,...,n) \\ \sum_{j=1}^{n} \hat{m}_j(\alpha_j,\beta_j,\lambda_j)x_j = r_G \\ \sum_{j=1}^{n} x_j = 1 \end{cases} \tag{8}$$

From the first equation in (8), we obtain the following solution:

$$x_j^* = \exp\left[\frac{1}{w}(\gamma\hat{m}_j(\alpha_j,\beta_j,\lambda_j) + \eta - (1-w)F(\tilde{\bar{r}}_j) - w)\right] \tag{9}$$

By substituting this solution in second and third equations in (8) and performing equivalent transformations, we obtain the equation system for determination of the optimal Lagrange multipliers:

$$\begin{cases} \eta^* = -\log\left(\sum_{j=1}^{n} \exp\left[\frac{1}{w}\left(\gamma^* \hat{m}_j\left(\alpha_j, \beta_j, \lambda_j\right) - (1-w) F\left(\tilde{\bar{r}}_j\right) - w\right)\right]\right) \\ \sum_{j=1}^{n} \left(\hat{m}_j\left(\alpha_j, \beta_j, \lambda_j\right) - r_G\right) \exp\left[\frac{1}{w}\left(\gamma^* \hat{m}_j\left(\alpha_j, \beta_j, \lambda_j\right) - (1-w) F\left(\tilde{\bar{r}}_j\right) - w\right)\right] = 0 \end{cases} \tag{10}$$

Consequently, by solving second equation in (10) on γ, we obtain the exact optimal portfolio of our proposed model (6). Furthermore, from optimal portfolio (9), the mathematical form is exponential, and so each portfolio rate is seldom 0. This means that our proposed model will achieve the diversification avoiding the risk-averse derived from the hybrid entropy of fuzzy random returns more than previous Xu's models.

4 Conclusion

The new entropy-based fuzzy random portfolio selection problem was proposed by extending previous Xu's models with the hybrid entropy. Our proposed model was initially formulated as a bi-objective programming problem, and in order to solve it directly in mathematical programming, the aggregated function to both objects were introduced. Then, the exact solution algorithm to obtain the optimal portfolio was developed using nonlinear programming approaches, particularly the KKT condition. From the exact exponential form of optimal portfolio, each portfolio rate is seldom 0, and so our proposed model will achieve the diversification avoiding the risk-averse derived from the hybrid entropy of fuzzy random returns.

As future studies, in order to obtain different and significant features compared with previous portfolio models or different types of investors' subjectivities, we need to perform an sensitivity analysis for optimistic-pessimistic parameters and weight of aggregated objective function using some practical market data.

Acknowledgements. This work was supported by The Ministry of Education, Culture, Sports, Science and Technology (MEXT), Grant-in-Aid for Young Scientists (B) (22700233).

References

1. Al-sharhan, S., Karray, S., Basir, O.: Fuzzy entropy: a brief survey. In: IEEE Proceedings of the 10th International Conference on Fuzzy Systems, vol. 3, pp. 1135–1138 (2001)
2. Bawa, V.S., Lindenberg, E.B.: Capital market equilibrium in a mean-lower partial moment framework. Journal of Financial Economics 5, 189–200 (1977)
3. Campos, L.M., Gonzalez, A.: A subjective approach for ranking fuzzy numbers. Fuzzy Sets and Systems 29, 143–153 (1989)
4. Carlsson, C., Fuller, R.: Fuzzy Reasoning in Decision Making and Optimization. Physica Verlag (2002)

5. Elton, E.J., Gruber, M.J.: Modern Portfolio Theory and Investment Analysis. Wiley, New York (1995)
6. Gil, M.A., Lopez-Diaz, M.: The λ-average value and the fuzzy expectation of a fuzzy random variable. Fuzzy Sets and Systems 99, 347–352 (1998)
7. Inuiguchi, M., Ramik, J.: Possibilisitc linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111, 3–28 (2000)
8. Katagiri, H., Ishii, H., Sakawa, M.: On fuzzy random linear knapsack problems. Central European Journal of Operations Research 12, 59–70 (2004)
9. Katagiri, H., Sakawa, M., Ishii, H.: A study on fuzzy random portfolio selection problems using possibility and necessity measures. Scientiae Mathematicae Japonocae 65, 361–369 (2005)
10. Konno, H., Yamazaki, H.: Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market. Management Science 37, 519–531 (1991)
11. Konno, H., Shirakawa, H., Yamazaki, H.: A mean-absolute deviation-skewness portfolio optimization model. Annals of Operations Research 45, 205–220 (1993)
12. Kwakernaak, H.: Fuzzy random variable-I. Information Sciences 15, 1–29 (1978)
13. Leon, R.T., Liern, V., Vercher, E.: Validity of infeasible portfolio selection problems: fuzzy approach. European Journal of Operational Researches 139, 178–189 (2002)
14. Markowitz, H.M.: Portfolio Selection. The Journal of Finance 7(1), 77–91 (1952)
15. Nieradka, G., Butkiewicz, B.S.: A Method for Automatic Membership Function Estimation Based on Fuzzy Measures. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS (LNAI), vol. 4529, pp. 451–460. Springer, Heidelberg (2007)
16. Philipatos, G.C., Wilson, C.J.: Entropy, market risk, and the selection of efficient portfolios. Applied Economics 4, 209–220 (1972)
17. Popkov, A.Y.: Entropy model of the investment portfolio. Automation and Remote Control 67(9), 1518–1528 (2006)
18. Puri, M.L., Ralescu, D.A.: Fuzzy random variables. Journal of Mathematical Analysis and Applications 114, 409–422 (1986)
19. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. Journal of Risk 2(3), 1–21 (2000)
20. Simonelli, M.R.: Indeterminacy in portfolio selection. European Journal of Operational Research 163, 170–176 (2005)
21. Smimou, K., Bector, C.R., Jacoby, G.: A subjective assessment of approximate probabilities with a portfolio application. Research in International Business and Finance 21, 134–160 (2007)
22. Tanaka, H., Guo, P.: Portfolio selection based on upper and lower exponential possibility distributions. European Journal of Operational Researches 114, 115–126 (1999)
23. Tanaka, H., Guo, P., Turksen, I.B.: Portfolio selection based on fuzzy probabilities and possibility distributions. Fuzzy Sets and Systems 111, 387–397 (2000)
24. Vercher, E., Bermudez, J.D., Segura, J.V.: Fuzzy portfolio optimization under downside risk measures. Fuzzy Sets and Systems 158, 769–782 (2007)
25. Watada, J.: Fuzzy portfolio selection and its applications to decision making. Tatra Mountains Math. Pub. 13, 219–248 (1997)
26. Xu, J., Zhou, X., DashWu, D.: Portfolio selection using λ mean and hybrid entropy. Annals of Operations Research 185, 213–229 (2011)
27. Yang, J., Qiu, W.: A measure of risk and a decision-making model based on expected utility and entropy. European Journal of Operational Research 164, 792–799 (2004)
28. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)

Formalization of Decentralized Cooperation Model for Disaster Simulation

Bipin Khanal, Nobuhiro Ito, Kazunori Iwata, and Naohiro Ishii

Abstract. In Disaster Simulation, multiple heterogeneous agents work together, forming a cooperative team, in a common environment to achieve a common goal. As it is difficult to achieve the expected results in dynamic environment, the current Disaster Simulations are forced to use centralized system in which the central agent is handling the communication, cooperation and coordination activities of all other agents. In centralized system, if the central agent (*that is controlling the whole system*) fails then the whole system could crash. Also, if we cannot skip the effect of dynamic environment in centralized system, it may result in more unwanted communication and, hence, may impair the performance of the whole system. This paper focuses on such performance bottlenecks caused by centralized model introducing a decentralized model for cooperation and its formalization. We describe how the decentralized model could be the best solution for robustness of the performance of the Disaster Simulation and we also discuss about how to formalize the group for decentralized cooperation model in dynamic environment.

1 Introduction

Disaster Simulation is a Multi-Agent System where multiple heterogeneous agents work together, forming a cooperative team, to achieve a common goal. A Multi-Agent System (MAS) is a system composed of multiple interacting intelligent agents. Multi-agent systems can be used to solve problems that are difficult or impossible for an individual agent or a monolithic system to solve [1]. Natural disasters are becoming very common these days. Disaster Simulation (hereafter DS) is an approach to rescue victims of disaster and reduce disaster damages.

Bipin Khanal · Nobuhiro Ito · Naohiro Ishii
Department of Information Science, Aichi Institute of Technology

Kazunori Iwata
Department of Business Administration, Aichi University, Aichi Ken, Japan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 205–214.
springerlink.com

In DS, participating agents such as the fire brigade, ambulance team, and police force work together to achieve the common goal of rescuing victims of disaster in a given environment [2]. In other words, Instead of human beings, robotic agents with different capacities play the role of rescuers and save people in the case of large-scale disaster.

1.1 Existing Concept of DS

Current DS is using centralized system in which the central agent is handling the communication, cooperation and coordination activities of all other agents. A number of researches related to the change in environment, i.e. dynamic environment in MAS are ubiquitous but the implementation of their results is still lacking. It is difficult to achieve the expected results in dynamic environment and, hence, the current DS is using the centralized system. The use of central system can slow down the performance of the whole system if the communication is increased. Specifically, having one agent in charge of handling the communication, cooperation and coordination of the activities of all other agents could result in a performance bottleneck [4]. In centralized system, if the central agent fails, then the whole system could crash.

1.2 Current Issues in DS

Like in MAS, all other things except agents are their environments in DS. According to Gemma Bel-Enguix and M. Dolores Jiménez-López (2007) [3], "*the environment has been recognized as an explicit and exploitable element to design Multi-Agent Systems (MAS). It can be assigned a number of responsibilities that would be more difficult to design with the sole notion of agents*". In other words, environment is a first-class abstraction that provides the surrounding conditions for agents to exist and that mediates both the interaction among agents and the access to resources. Any changes in the environment can have enormous effect in the movement of agents in DS. This is one of the main issues in DS. If we skip the effect of dynamic environment, then the use of centralized system could be the best approach. But when we have to consider the dynamic environment, the messages related to the environment may increase and the use of centralized system could arise a question mark in the performance of the whole system.

Only considering the dynamic environment will not solve the problem of current DS. Modeling the cooperation of agents in a dynamic environment is of utmost importance. In this paper, we mainly focus on these issues. To make the system decentralized, we need to re-structure the DS and re-model the cooperation of the agents.

2 Environment as an Agent

According to Nobuhiro Ito, Kouichi Nakagawa, Takahiro Hotta, Xiaoyong Du, Naohiro Ishii (1998) [6], we got a clear concept of modeling environment as an

agent in MAS. In this research, we consider environment as an agent because of the following reasons.

1. *To decrease the unnecessary communication by acting as a message filter: Environment agent* receives all the messages and only necessary messages will be sent to agents within the very environment after filtering the unnecessary messages.
2. To allow agents to calculate independently: Since the *Environment agent* sends all the required information to all agents, they can make calculations independently irrespective of other agents.
3. To provide a medium for data transfer for all agents within its area: Communications between agents are done via *Environment agent*.
4. To take part in cooperation, so that an effective cooperative team could be generated: If *Environment agent* also takes part in cooperation, a team including the *Environment agent* can be created and thus the updated information can be tracked by all agents when necessary so that the cooperation is not violated unnecessarily. Considering Environment as an agent, the system will be decentralized providing independent calculation and communications among all the agents. Now we discuss the new approach of the decentralized disaster simulation.

2.1 A New Structure of DS

We consider Environment as one of our agents in our new structure of DS. The figure below describes the clear view of our new structure.

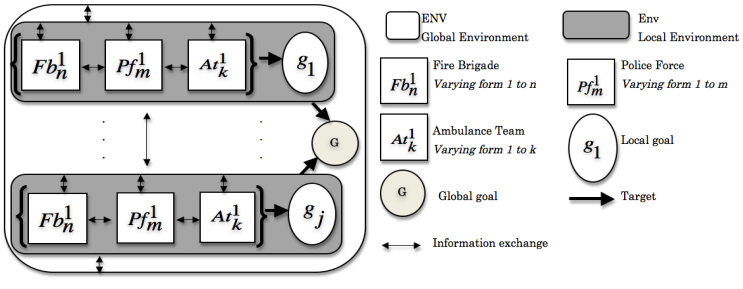


Fig. 1 Structural diagram showing relations between agents

In Fig.1, we have a Global Environment (*ENV*) that contains sub environments and a common goal (*G*). Sub environment here is considered as Local environment (*Env*) that includes Fire brigade (*Fb: varying in number from 1 to n*), Police force (*Pf: varying in number from 1 to m*), Ambulance team (*At: varying in number from 1 to k*) and Buildings as local goals (*g: varying in number from 1 to j).* All the agents *Fb, Pf and At* work together to achieve a local goal (*g)* within a Local Environment (*Env)* provided that the number of all the agents taking part in

the action may vary correspondingly. Once all the local goals (*g=number of buildings: varying from 1 to j)* are accomplished, then global goal G is supposed to be accomplished.

2.2 A New Approach to DS

In this approach, we consider environment as an active agent that has a big role in DS. We are introducing some additional attributes to the agents like *Status, Priority and Free Flag*, which are of utmost necessity while working in a dynamic environment.

Status S: It describes the current environmental status as compared to the previous one and is defined by using 0 or 1. Initially it is 0, if there is a change in environment, it is set to 1 and it is 0 for next change and so on.

Priority P: It is describes the priority of each agent. Priority can be categorized into three levels.

1. *High (2)*: Agents that have a very active role have high priority. Initially all agents have high priority.
2. *Medium (1)*: Agents whose role is not very active have this priority. For example, in the case of heavy rain, the use of Fire Brigade can be either suppressed or avoided. These types of agents are considered to have a medium priority.
3. *Low (0)*: Agents that are out of order or that do not have active role in the team.

From above descriptions, the attribute *Priority* is very important in case of dynamic environment. The Priority is necessary to avoid the unnecessary involvements of disturbing agents in the cooperation. Agents enter the cooperation on the basis of priority.

Free Flag F: This attribute defines whether the agent is free or not. For free agents, it is set to 1 and 0 is set for busy agents.

The step-wise explanation of our new approach is described as follows.

STEP 1: Global Environment *(ENV)* sends data to all other agents by setting the attributes S to 0, P to 2, and F to 1.

```
send_to_all_agents(S,P,F,…)
```

STEP 2: As all agents know their position, they can calculate their distances to the buildings in a given environment and send necessary data to all other agents. The data includes *myID(my_id), BuildingID(Bld_id), Distance(D),Status(S), Priority(P), Free Flag(F) etc.*

```
receive_from_ENV(S,P,F,…)
calculate_Distance()
send_to_Agents(my_id, Bld_id, D,S, P, F,…)
```

STEP 3: The agents always accept messages from the Environment and messages from other agents only when they are free (initially all agents accepts messages from each other). All agents create their own task list using the data sent by other agents based on the priority and the shortest distance. As a result, the agent with the shortest distance to the building with high priority comes on the top of the list and so on. Virtual agent called Local Environment (*Env*) is created and task list is updated to it. *Env* creates a group for cooperation and sends signal to agents to start cooperation and work by setting the F to 0 for all agents within it.

A simple example of program flow of agent when it is free can be described as follows:

```
if(F==1){
   if(msg_from_agents){
        receive_From_Agents(my_id, Bld_id, D,S, P, F,…)
          create_task_list()
          update_Task_List_To_Env()
       }else if(msg_from_Env){
          receive_msg_of_Env()
          cooperate()
         if(finished){
          set F=1
          send_to_Env(finished_task,F,…)
         }else{
          work()
         }
    }
}
```

Once the task is accomplished, F is set to 1 so that agents can take part in another task.

STEP 4: After the completion of the task in the task list, *Env* updates the accomplished task to *ENV* and all the above steps from STEP 2 are repeated recursively unless all the local goals are accomplished.
If there is change in *Env*, it re-allocates the task to the agents within it without disturbing the other *Env* agents. But if *ENV* is changed, all the processes are repeated from STEP1.

3 Group Formation

The group formation begins when the initiator recognizes the potential for cooperative action in order to reach the goal. In our research, we consider *Env* as an initiator. The stage of potential recognition is concerned with finding the set of agents that may participate in the formation of the group aiming to realize the common goal [8].

In our model, the members with same local goal g ($g \subseteq G$) can form a group if P = 2 or 1 (for effective group P=2) and F=1. The group formation takes place after the task list is updated to *Env*. The members of the group should not only be willing to cooperate with others but they should also have relevant skills and resources at their disposal.

3.1 Role of Env in Group Formation

The first task of local *Env* is to form a partial plan for achieving the local goal g. According to the goal, it determines the characteristics of agents that are most suitable to form the group. In order to determine this match, the *Env* needs to find out relevant properties of the agents, being interested in four aspects: their *abilities*, *willingness*, *opportunities* and *individual type*.

***Ability and willingness*:** These characteristics of agents can be decided by their Priority P. If P=2 or 1, we can consider that agents have willingness to and are able to achieve the goal.

able(RA,g) iff P=2 or P=1 where $RA = \{AT, PF, FB, Env\}$
willing(RA,g) iff P=2 or P=1 where $RA = \{AT, PF, FB, Env\}$

***Opportunities*:** This property of agent can be defined by its Free flag F. If F=1, agent can be considered to have opportunity to enter the group.
opp(RA,g) iff F=1 where $RA = \{AT, PF, FB, Env\}$

***Individual type*:** The group could be effective if agents with different types participate in it. All agents have different capabilities. To form an effective group *Env* chooses agents with different individual types. For example, the participants for the group will be FB PF, AT and Env. There should be proper type of distribution of agents in the group to achieve the goal g and that is denoted by

indvtype(At, g) where $At \subseteq AT$
indvtype(Pf, g) where $Pf \subseteq PF$
indvtype (Fb, g) where $Fb \subseteq FB$

With the help of above defined properties, the *Env* can decide whether the agent can form a group or not.

$$can(At, g) \leftrightarrow able(At, g) \wedge opp(At, g) \wedge willing(At, g) \wedge indvtype(At, g)$$
$$Group(Ra, Env) \leftrightarrow can(At_t, Env_t) \wedge can(Pf_t, Env_t) \wedge can(Fb_t, Env_t)$$

where t is the time step of the simulation.

4 A Cooperation Model

A cooperative system is defined to be a system of multiple dynamic entities that share information or tasks to accomplish a common, though perhaps not singular, objective [7]. In this paper, we consider agents as our multiple dynamic entities that share information to achieve a common target/goal. For cooperation, the most important point is communication. Wrong information may lead to the wrong result. So there should be proper message sharing for good cooperation. In this paper, we suppose that there is a very good message-sharing environment between

the agents. In DS, all the agents have a common goal of rescuing victims of disaster.

In this section, we model the cooperation for the agents.

$AT = \{At_1, At_2, ..., At_n\}$, $PF = \{Pf_1, Pf_2, ..., Pf_m\}$, $FB = \{Fb_1, Fb_2, ..., Fb_k\}$,
$ENV = \{Env_1, Env_2, ..., Env_j\}$

The sets written above are possible sets of Ambulance Team, Police Force, Fire Brigade and Environment respectively. In our model, we use '*t*' as one of the time steps of simulation.

The rescue team can be expressed as $RA = \{AT, PF, FB, Env\}$

In DS, we can consider the building as a goal for agents.

$$G = \{Bld_1, Bld_2, ..., Bld_j\}$$

Here the number of buildings (local goals) and virtual Local Environment agents are same because each goal is inside a virtual environment agent.

Common objective is defined as a finite set of global goals.

$$O_c = \{G_1, G_2, ..., G_m \mid G_i \in G, m \geq 1\}$$

In disaster simulation, since the common objective of all the agents is only one, i.e. to rescue, m is 1.

Model of cooperation can be related to the common objective as below:

$$Mod \overset{def}{=} (Ra, O_c) \text{ where } Ra \subseteq RA$$

Relation between agent and common objective can be expressed as $\text{Re}l\_Agent : RA \times O_c \rightarrow P(LG)$ where $LG \subseteq G$ and *P(LG)* = Possible Local Goals.

Cooperation Capability can be defined as a set of services and functionalities needed to support cooperation, such as negotiation, collaboration and coordination which can be expressed as below.

$$CF \overset{def}{=} \{AT_{cf}, PF_{cf}, FB_{cf}, ENV_{cf}\}$$

We represent service or control functionality by a function that brings the system from a control state to another. CS is the set of Control state.

$$cf_i : CS \rightarrow CS$$

$$cs_t^i \ \mathsf{a} \ cs_{t+1}^i$$

$$CC : RA \rightarrow P(CF)$$

The function *Cooperation* measures the ability of a society of agents to achieve the cooperation control of a global goal.

$$Cooperation : P(RA) \times G_g \rightarrow \{true, false\} \; where \; G_g \subseteq O_c$$

An agent is able to support the cooperation needed to reach a goal Bld_1 if its cooperation capability allows bringing the system from the initial control state to a state of describing a goal.

$$Cooperation(At_1, Bld_1) = true \; iff \; \exists Exec \in (Cap(At_1))^*.Exec(s_i) = State(Bld_1)$$

$$[f_1, ..., f_n](s_i) = f_n(...(f_1(s_i)))...)$$

The functionalities here are only concerned with performing domain task.

Completeness of local goals:
Completeness of local goal can be expressed as below.

$$Comp(RA, G) = true \; iff \; \{Bld_1, Bld_2, ..., Bld_3\} = \bigcup_{t=1}^{n} \mathrm{Re}\, l\_Agnet(A_t, G)$$

$$where \; A_t = \{At_t, Pf_t, Fb_t, Env_t\}$$

Efficiency of local goals:
It states the fact that only one agent will realize the local goal.

$$Effic(RA, G) = true \; iff \; \forall i, j . i \neq j \Rightarrow \mathrm{Re}\, l\_Agent(A_i, G) \cap \mathrm{Re}\, l\_Agent(A_j, G) = \{\}$$

Consistency of local goals:
It states that no agent of the society can destroy the results achieved by another agent.

$$Cons(RA, G) = true \; iff$$

$$\forall i, j. \; \mathrm{Re}\, l\_Agent(A_i, G) \subset (\mathrm{Re}\, l\_Agent(A_i, G) \oplus \mathrm{Re}\, l\_Agent(A_j, G))$$

The operation $\oplus$ is defined on the set of local goals as follows: if the result of an agent is the opposite of the result achieved by another agent then the addition ($\oplus$) of their local goals corresponds to the neutral goal represented by the empty set.

For the proper implementation of the cooperation in DS, we need to consider the following points.

Role: The notion of role allows identify and specifying a particular agent. It also allows characterizing the responsibilities associated to each agent.

$$Role : RA \times P(O_c) \rightarrow P(LG)$$

$$(A, Bld_1) \; \mathrm{a} \; U_{B_{gk} \in O_c} \mathrm{Re}\, l\_Agent(A, B_{gk})$$

Mode of Cooperation: It regroups the role of cooperative agents. Let $Mod = (\{At_t, Pf_t, Fb_t, Env_t\}, O_c)$ where t is the time step of the simulation. A mode of cooperation is represented by n-tuple of roles.

$$m \overset{def}{=} (r_1, r_2, ..., r_n) \text{ and } M(\mathrm{mod}) \overset{def}{=} \{(r_1, ..., r_n) \mid r_i \in Role(A_i, O_c), i \in \{1, ..., n\}\} \text{ where}$$

$$A_t = \{At_t, Pf_t, Fb_t, Env_t\}$$

A mode of cooperation is said to be optimal if same goal will not achieved by two different agents within the same mode and can be expressed as $\forall i, j . r_i \cap r_j = \varnothing$.

***A cooperation strategy*:** A strategy of cooperation is a process that controls the well functioning of the mode of appropriate cooperation in a given situation. This depends on the mode of cooperation environment.

Env_{sys} be the set of environmental states; then $E_{sys} = \{Env_1, Env_2, ..., Env_j\}$

We can define cooperation strategy as

$$ST : M \times E_{sys} \rightarrow M$$

$$(m, Env_i) \text{ a } m'$$

$U_s(m, E_{sys})$ describes the states of agents and environment. A model m is adopted in Env_i of the system if and only if the condition $U_s(m, Env_i)$ satisfied.

$$ST(m, Env_i) = m' \; iff U_s(m, Env_i, m') = true$$

5 Implementation and Evaluation

We implemented our idea in the current Robocup Rescue Simulation and found that our Environment agent, as a mediator and message filter, helped in effective group formation for efficient cooperation, which was quite difficult in former simulation. Fig. 2 and Fig. 3 as below show the clear concept of message filtering using Environment agent.

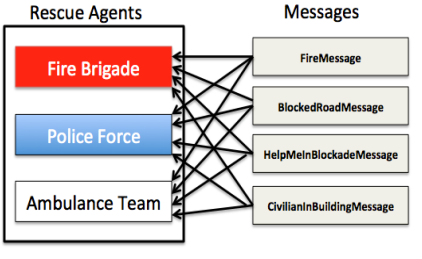


Fig. 2 Before Implementation

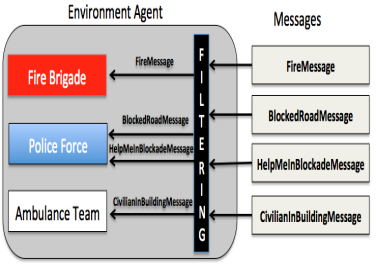


Fig. 3 After Implementation

We were also able to prove that a change in one environment/group does not affect the agents' action in any other environment/group. With the implementation of our concept, agents could decide their own goal and hence, they could work until the very goal was accomplished without any unnecessary disturbances. We tested our idea by increasing the number of agents in the simulation and found effective groupism and communication between agents with the least possible message sharing by neglecting unnecessary messages. Our result also satisfied the

condition that "*if control and responsibilities are sufficiently shared among different agents, the system can tolerate the failure of one or more of the agents resulting in effective performance*."

6 Conclusions and Future work

In this research, we described about the issues in current DS and concluded that the change in environmental condition is one of the main issues. To solve the problem, we introduced a new approach of decentralized model and described how the decentralized model could be the best solution for effective performance of the DS in comparison with the centralized one. The benefits of decentralized system over centralized one are as under.

1. Calculations are equally divided to all agents,
2. Agents can work independently irrespective of the action of other agents.
3. All agents equally share communication, cooperation and coordination.
4. Even if one part of the system crashes, other parts can be working properly.

To implement the effect of the dynamic environment, we modeled environment as an agent and re-structured the current DS. We described that the performance of the whole system could be more effective if we use the decentralized approach considering the environment as an agent. The environment agent helped us to formalize a group and our cooperation model helped us to make a clear view of good cooperation between the agents in a dynamic environment.

References

1. http://en.wikipedia.org/wiki/Multi-agent_system
2. Khanal, B., Ito, N., Syamoto, T., Iwata, K.: RoboCup Rescue 2011–Rescue Simulation League Team Description Paper<Naito Rescue 2011> (2011)
3. Bel-Enguix, G., Jiménez-López, M.D.: Agent-Environment Interaction in a Multi-agent System: A Formal Model. In: Genetic and Evolutionary Computation Conference (2007)
4. Rahwan, T.: Algorithms for Coalition Formation in Multi-Agent Systems: A PHD thesis, University of Southampthon (2007)
5. Jmaiel, M., Kacem, H.: A Formal Model for Cooperation and Agency in Multi-Agent System, University of Sfax LARIS Laboratory (1999)
6. Ito, N., Nakagawa, K., Hotta, T., Du, X., Ishii, N.: EAMMO: an environmental agent model for multiple objects. Information and Software Technology 40 (1998)
7. http://www.springer.com/series/5788
8. Duinin-Keplicz, B., Verbrugge, R.: Teamwork in Multi-Agent systems. A Formal Approach (2010)

Fuzzy Decision Making Model for Human Fall Detection and Inactivity Monitoring

Marina V. Sokolova and Antonio Fernández-Caballero

Abstract. In this paper a fuzzy decision making model for fall detection and inactivity monitoring is presented. The inputs to the model are regions of interest (ROIs) corresponding to humans segmented in infrared video. The classification features proposed in fuzzy decision include geometric and kinematic parameters associated with more or less sudden changes in the tracked bounding boxes of the ROIs. The paper introduces a complete fuzzy-based fall detection system capable of identifying true and false falls, enhanced with inactivity monitoring aimed at confirming the need for medical assistance and/or care. The fall indicators used as well as their fuzzy model is explained in detail. As the proposed model has been tested for a wide number of static and dynamic falls, some exciting initial results are presented here.

1 Introduction

Fall detection is to date a challenging problem [1], especially for monitoring elderly people [5, 10]. The process of efficient fall detection can be seen as complex decision making. Moreover, fuzzy decision making is a powerful paradigm for dealing with human expert knowledge when designing fuzzy-based models. Starting from segmented video sequences, several approaches to knowledge-based fall detection methods have been proposed so far. For instance, head motion trajectory analysis is a major indicator of a fall [18]. In [22] the authors propose a 3D method from 2D

Marina V. Sokolova
Instituto de Investigación en Informática de Albacete (I3A),
Universidad de Castilla-La Mancha, 02071-Albacete, Spain,
and South-West State University, 94, 50 let Oktyabrya, 305040 - Kursk, Russia
e-mail: Marina.Sokolova@uclm.es

Antonio Fernández-Caballero
Departamento de Sistemas Informáticos, Universidad de Castilla-La Mancha,
02071-Albacete, Spain
e-mail: Antonio.Fdez@uclm.es

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 215–224.
springerlink.com

pose reconstruction based on the introduction of bone, bone symmetry, and rigid body projection constraints. Another approach for fall detection, based on the *k*-nearest neighbor algorithm, is introduced in [12]. Moreover, 3D shape-encoded filters are described in [15]. These extract the boundary gradient information of the body image, which is then propagated by a branching particle system. In [19], the application of a kernel particle filter for 3D body tracking is introduced. In [14] a system for human behavior analysis is reported, where a human is represented with three components: human star skeleton, angles of six sticks in the star skeleton, and object motion vectors.

This paper starts from proper recent advances in infrared human detection [8, 2, 9]. A natural objective after efficient human segmentation in infrared video is to get powerful human fall detection models. In infrared segmentation, consider, for instance, a pedestrian detection method in thermal infrared imagery using the wavelet transform, as proposed in [11]. In [4], a background subtraction is initially performed to identify local foreground objects. Then, gradient information from the objects and the background is combined to elaborate a contour map representing the confidence for each pixel to belong to a human's boundary. Finally, a path search is performed to complete broken segments. An approach based on histogram is introduced on [7], realizing a vertical and a horizontal adjustment on the pedestrian candidate regions. More recently, the Otsu algorithm [17] is used to binarize the image separating the warmer regions (probably belonging to humans) from the colder ones, using opening and closing morphological operations to remove small broken parts obtained from the binarization [3]. Next, objects that are too small or have a wrong aspect ratio to be humans are filtered out and a histogram of oriented gradients (HOG) method is used to extract features used to train an Adaboost classifier.

In this paper, we introduce a new human fall and inactivity detection model by using fuzzy decision making. Indeed, fuzziness has been largely studied in the past for pattern recognition [6, 21]. In relation to fall detection, a method with fuzzy one class support vector machine in a smart room has been presented in [23]. The fuzzy membership function represents the likelihood for the data to belong to a target class [16]. This is also our approach, as we consider that fuzzy-based modeling addresses a more flexible decision making process. The paper is organized as follows. Section 2 introduces the proposed approach towards human monitoring and fall detection, explaining our fall detection and inactivity monitoring subsystems. Then, in Section 3 data and experimental results are presented. Finally, some final conclusions are provided in Section 4.

2 Fall Detection and Inactivity Monitoring System

Our approach is based on the idea of combining a fall detection with an inactivity monitoring model. In this case, the inactivity monitoring subsystem is thought for checking if a human continues laying inactive during some predefined time after a fall has been detected in the detection subsystem. In this paper, fall detection as such is based on the geometrical analysis of the segmented region of interest (ROI)

representing a human. Roughly, it is decided that a fall has occurred through the study of the velocity of the deformation suffered by the ROI along an established time interval. This time interval is denominated the "fall time", t_f, which typically takes values between 1 and 3 seconds. Additionally, fuzzy logic is used in order to make decisions more flexible, to avoid some limitations in the representation of human figures, and to smooth the limits of the evaluation parameters. The objectives of our fall detection approach are manifold, as they include (a) to detect falls (severe and non-severe), (b) to identify false falls, such as kneeling and bending, and, (c) to claim for assistance in case of severe falls.

Fig. 1 illustrates the principal ideas of our approach towards fall detection and inactivity monitoring. In brief, during fall time, f_t, the fall detection subsystem is launched ("A" in Fig. 1). Once a fall is detected, the inactivity monitoring subsystem is activated ("B" and "C" in Fig. 1). If the fallen human stands up before a pre-established time interval, denominated "monitoring time", t_m (typically 30 seconds), the inactivity monitoring subsystem is stopped, and the fall detection subsystem is launched again ("D" and "E" in Fig. 1).

2.1 Fall Detection Subsystem

Firstly, let us explain in more detail the proposed human fall detection subsystem. In our proposal, the general idea of fall detection is based on the joint static and kinematic analysis of the ROIs corresponding to humans. The statical analysis corresponds to the study of a single (and current input) segmented infrared image, whilst kinematic analysis is based on a sequence of segmented infrared video images. During statical analysis, the geometrical parameters of each present ROI are calculated, namely, the "change of the height of the ROI", Δh, and the "width-to-height ratio of the ROI", R_{wh}. The kinematic analysis implies the calculation of the "velocity of the change of the height of the ROI", Δv, among a series of n consecutive segmented video frames. All these parameters, together with the predefined "fall time", t_f, compose a set of fall indicators used in this research.

2.1.1 Calculation of the Fall Indicators

For the n consecutive ROIs corresponding to a same human, let $y_{tr,1}$ and $y_{dl,1}$ be the top-right and down-left Y-coordinates of the ROI in the first image, and let $y_{tr,n}$ and $y_{dl,n}$ be the top-right and down-left corner Y-coordinates of the ROI in the last (n-th) image , respectively. In the same sense, let $x_{tr,1}$ and $x_{dl,1}$ be the top-right and down-left X-coordinates of the ROI in the first image, and let $x_{tr,n}$ and $x_{dl,n}$ be the top-right and down-left corner X-coordinates of the ROI in the n-th image. Let us also define x_{tr}, x_{dl}, y_{tr} and y_{ld} as the top-right and down-left coordinates of the current ROI. Also, let $w = x_{tr} - x_{dl}$ and $h = x_{tr} - x_{dl}$ be the width and the height, respectively, of the current ROI. The parameters are calculated as follows:

$$R_{wh} = \frac{x_{tr} - x_{dl}}{y_{tr} - y_{dl}} \tag{1}$$

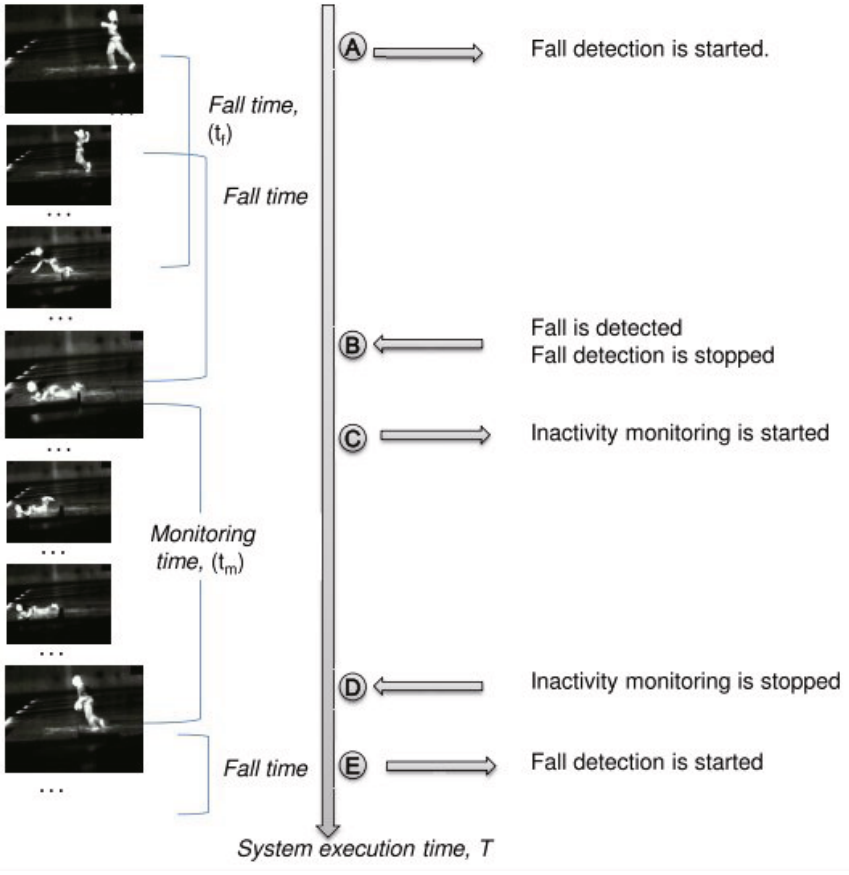


Fig. 1 Graphical representation of the fall detection and inactivity monitoring system

$$\Delta h = \frac{y_{tr,1} - y_{dl,1}}{y_{tr,n} - y_{dl,n}} \tag{2}$$

$$\Delta v = \frac{\sqrt{[(x_{tr,1} - x_{dl,1}) - (x_{tr,n} - x_{dl,n})]^2 + [(y_{tr,n} - y_{dl,n}) - (y_{tr,n} - y_{dl,n})]^2}}{\sqrt{w^2 + h^2}} \tag{3}$$

Notice that the denominator in formula (3) is the maximum possible velocity of a fall.

2.1.2 Fuzzy Model for the Fall Indicators

After the previous parameters have been calculated, fuzzy logic is used in the proposed fall detection subsystem to determine the ranges for the fall indicators.

Fig. 2 From left to right, representation of the linguistic variables "HeightChange", "VelocityChange" and "WidthToHeightRatio" in Spark! Fuzzy Logic Editor

Indeed, fuzziness for fall detection includes some input linguistic variables, namely "HeightChange", "VelocityChange" and "WidthToHeightRatio", directly related to Δh, Δv and R_{wh}, respectively, and an output linguistic variable called "Fall". The upper and lower bounds for the values of the fuzzy sets of the linguistic variables are real numbers within the intervals provided:

```
VelocityChange   REAL; (* RANGE(0 .. 1.25) *)
HeightChange     REAL; (* RANGE(0 .. 4) *)
WidthToHeightRatio  REAL; (* RANGE(0 .. 6) *)
```

The fuzzy sets for each of these input linguistic variables include three fuzzy terms, namely SMALL, MEDIUM and HIGH. Fig. 2 shows the linguistic variables within the Spark! Fuzzy Logic Editor [20].

The linguistic variable "VelocityChange" is calculated as the relation of the real velocity to the maximum possible velocity through formula (3). Thus, in case of a very quick fall with velocity equal or close to the maximum, the following expression $\Delta v \approx 1.25$ is gotten. On the contrary, the change of velocity is minimal when $\Delta v \approx 0.0$. The terms for this linguistic variable have the following ranges:

SMALL$\in [0 \ldots 0.35]$
MEDIUM$\in [0.25 \ldots 0.65]$
HIGH$\in [0.6 \ldots 1.25]$

The linguistic variable "HeightChange" is calculated as the relation of the initial height to the final height of the ROI. As stated in [13], the width of a human's body is around 25 percent of his/her height. Taking into account different human builds, we have experimentally decided set up a value of 30 percent of a ROI's height as the higher boundary for the HIGH term of linguistic variable "HeightChange".

The "WidthToHeightRatio" variable ranges from 0.0 to 6.0. Since a ROI corresponding to a person in a strengthened position is longer in height than in width, the term SMALL $\in [0 \ldots 0.8]$, the term MEDIUM $\in [0.7 \ldots 2.5]$ stands for a situation where a ROI belongs to a fallen person, or, alternatively, to a sitting or prone person. Finally, the term HIGH $\in [2.1 \ldots 6.0]$ describes a ROI of a laying person.

The complete set of fuzzy rules used in this proposal to determine if a fall has occurred ("Fall is yes") is as follows:

```
RULE 0:  IF WidthToHeightRatio is HIGH AND HeightChange is HIGH
         AND VelocityChange is HIGH THEN Fall is yes;
RULE 1:  IF WidthToHeightRatio is HIGH AND HeightChange is MEDIUM
         AND VelocityChange is HIGH THEN Fall is yes;
RULE 2:  IF WidthToHeightRatio is HIGH AND HeightChange is HIGH
         AND VelocityChange is MEDIUM THEN Fall is yes;
RULE 3:  IF WidthToHeightRatio is HIGH AND HeightChange is SMALL
         AND VelocityChange is HIGH THEN Fall is yes;
RULE 4:  IF WidthToHeightRatio is HIGH AND HeightChange is HIGH
         AND VelocityChange is SMALL THEN  Fall is yes;
RULE 5:  IF WidthToHeightRatio is HIGH AND HeightChange is SMALL
         AND VelocityChange is MEDIUM THEN Fall is yes;
RULE 6:  IF WidthToHeightRatio is HIGH AND HeightChange is MEDIUM
         AND VelocityChange is SMALL THEN Fall is yes;
RULE 7:  IF WidthToHeightRatio is HIGH AND HeightChange is SMALL
         AND VelocityChange is SMALL THEN Fall is no;
RULE 8:  IF WidthToHeightRatio is SMALL AND HeightChange is HIGH
         AND VelocityChange is HIGH THEN Fall is no;
RULE 9:  IF WidthToHeightRatio is SMALL AND HeightChange is MEDIUM
         AND VelocityChange is HIGH THEN Fall is no;
RULE 10: IF WidthToHeightRatio is SMALL AND HeightChange is HIGH
         AND VelocityChange is MEDIUM THEN Fall is no;
RULE 11: IF WidthToHeightRatio is SMALL AND HeightChange is SMALL
         AND VelocityChange is HIGH THEN Fall is no;
RULE 12: IF WidthToHeightRatio is SMALL AND HeightChange is HIGH
         AND VelocityChange is SMALL THEN Fall is no;
RULE 13: IF WidthToHeightRatio is SMALL AND HeightChange is SMALL
         AND VelocityChange is MEDIUM THEN Fall is no;
RULE 14: IF WidthToHeightRatio is SMALL AND HeightChange is MEDIUM
         AND VelocityChange is SMALL THEN Fall is no;
RULE 15: IF WidthToHeightRatio is SMALL AND HeightChange is SMALL
         AND VelocityChange is SMALL THEN Fall is no;
RULE 16: IF WidthToHeightRatio is MEDIUM AND HeightChange is HIGH
         AND VelocityChange is HIGH THEN Fall is yes;
RULE 17: IF WidthToHeightRatio is MEDIUM AND HeightChange is MEDIUM
         AND VelocityChange is HIGH THEN Fall is yes;
RULE 18: IF WidthToHeightRatio is MEDIUM AND HeightChange is HIGH
         AND VelocityChange is MEDIUM THEN Fall is yes;
RULE 19: IF WidthToHeightRatio is MEDIUM AND HeightChange is SMALL
         AND VelocityChange is HIGH THEN Fall is no;
RULE 20: IF WidthToHeightRatio is MEDIUM AND HeightChange is HIGH
         AND VelocityChange is SMALL THEN Fall is no;
RULE 21: IF WidthToHeightRatio is MEDIUM AND HeightChange is SMALL
         AND VelocityChange is MEDIUM THEN Fall is no;
RULE 22: IF WidthToHeightRatio is MEDIUM AND HeightChange is MEDIUM
         AND VelocityChange is SMALL THEN Fall is no;
RULE 23: IF WidthToHeightRatio is MEDIUM AND HeightChange is SMALL
         AND VelocityChange is SMALL THEN Fall is no;
RULE 24: IF WidthToHeightRatio is MEDIUM AND HeightChange is MEDIUM
         AND VelocityChange is MEDIUM THEN Fall is yes;
```

2.2 Inactivity Monitoring Subsystem

Notice that there are two main variables for final decision making. The first one is *Alarm*, which is used to indicate if a fall has been detected. The variable holds values "FallDetected" and "NoFallDetected". Initially *Alarm* is set to "NoFallDetected" until a fall is supposed in accordance with the fall detection subsystem described before.

The other important variable is *Assistance*, indicating if there is a need to assist a fallen person. The possible values for *Assistance* are "AssistanceRequired" and "NoAssistanceRequired", being initially provided value "NoAssistanceRequired".

Table 1 Fall detection results

WidthToHeightRatio R_{wh}	HeightChange Δh	VelocityChange δv	*Alarm*	*Assistance*
1. True fall - Frontal fall				
MEDIUM (1.14)	HIGH (2.72)	MEDIUM (0.35)	FallDetected	AssistanceRequired
MEDIUM (1.53)	HIGH (3.04)	MEDIUM (0.41)	FallDetected	AssistanceRequired
MEDIUM (1.33)	MEDIUM (2.10)	SMALL (0.28)	NoFallDetected	NoAssistanceRequired
2. False fall - Kneeling person				
MEDIUM (1.35)	MEDIUM (1.37)	SMALL (0.16)	NoFallDetected	NoAssistanceRequired

Assistance is only valid after a fall has been detected and the inactivity monitoring mode has been launched. Thus, at the beginning of each fall time cycle, *Alarm* and *Assistance* take values "NoFallDetected" and "NoAssistanceRequired", respectively. The inactivity monitoring subsystem behavior may be described as follows:

1. When the fuzzy model detects a fall, variables *Alarm* and *Assistance* are set to "FallDetected" and "NoAssistanceRequired", which means that a fall has been detected and the inactivity monitoring is started. During inactivity monitoring, with arrival of the new ROIs, the Δh, R_{wh}, and Δv are calculated and then checked with the same fuzzy model.
2. If a person has fallen (that is, *Alarm* has value "FallDetected") and inactivity monitoring has been active for time interval t_m, *Assistance* is provided the value "AssistanceRequired". This means that medical care or assistance is needed.
3. In case the fuzzy model newly generates "NoFallDetected" for the *Alarm* variable, it is supposed that the fallen human has risen on his/her feet and the *Assistance* variable is set to "NoAssistanceRequired". In this case, the inactivity monitoring finishes and the fall detection subsystem is re-activated.

3 Data and Results

The experiments described next were set up in order to validate the proposed approach. The experiments were carried out on a personal computer equipped with an Intel Core i7 processor with 3 GB of RAM. The video sequences were recorded with a FLIR A320 infrared camera at a resolution of 720×480 pixels. Five series containing a single person were used to validate our research. The experiments were carried out in order to detect real and false falls. The results are offered in Table 1. Falls were simulated by an actor. The first column contains the number of the detection experiment. The following columns contain values for the calculated parameters Δh, R_{wh}, and Δv for each time fall period t_f. The *Alarm* column shows the response of the fuzzy model for the fall detection, which is "FallDetected" in case of a fall and "NoFallDetected" otherwise. *Assistance* is set to "AssistanceRequired" in case of inactivity after a fall and "NoAssistanceRequired" in case a person has risen.

Fig. 3 A frontal fall (captured from a lateral view)

Fig. 4 Identification of a false fall (kneeling person)

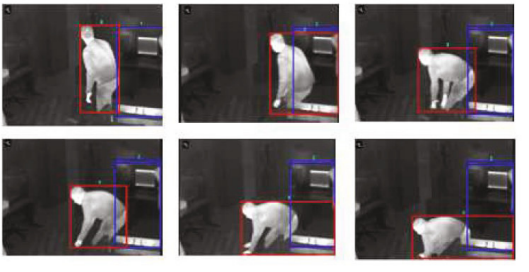


Fig. 3 and Fig. 4 show resulting images where falls are initially supposed in the experiment. There are red and blue bounding boxes in the images. A red bounding box shows the rectangular borders of the ROI of the last image, whereas a blue bounding box retains the borders corresponding to the first image originating the fall presumption. Indeed, fall indicators are calculated using the red and the blue bounding boxes. The capture of the video images was performed with an interval of 200 milliseconds, and the fall time was set to $t_f = 1.2$ seconds. Thus, fall detection is tested on every six consecutive frames. Table 1 contains the values of the evaluation parameters for these falls.

The first fall, denominated "frontal fall (lateral view)" (see Fig. 3), has the following characteristics at the moment of fall detection. "WidthToHeightRatio" is MEDIUM (from $R_{wh} = 1.14$), "HeightChange" is HIGH (from $\Delta h = 2.72$), and "VelocityChange" is MEDIUM (from $\Delta v = 0.35$). *Alarm* is "FallDetected" and *Assistance* is set to "AssistanceRequired". The same parameters calculated for the next image keep the same terms for their corresponding linguistic variables, that is MEDIUM (1.53), HIGH (3.04) and MEDIUM (0.41). Next, "VelocityChange" reduces to SMALL (0.28), and "HeightChange" decreases to MEDIUM (2.10) terms, which makes change values of variables *Alarm* and *Assistance* to "NoFallDetected" and "NoAssistanceRequired". See the three situations described on top of Table 1.

False falls are represented in this case study with the kneeling action. In fact, a false fall may be detected as a fall which lasts more than a fall time as well as having small values for fall detection parameters given in the last case shown in Table 1. As a rule, in case of a false fall, linguistic variable "WidthToHeightRatio" may have SMALL or MEDIUM values, and linguistic variable "VelocityChange" is rather SMALL. Fig. 4 shows a moment of a false fall. In compliance with the results provided in Table 1, a fall is not detected in this case. Thus, the algorithms proposed were able to differentiate true falls from false falls. Table 1 contains values for evaluation parameters for the "kneeling person" situation with 'WidthToHeightRatio" in MEDIUM (1.35), "HeightChange" in MEDIUM (1.37), and "VelocityChange" in SMALL (0.16).

4 Conclusions

Falls of elderly people is a big problem which needs rapid and effective solutions. In this paper we have presented a fuzzy-based model for fall detection and inactivity monitoring in previously segmented infrared video. The data supplied by a generic human detection algorithm are used by the fuzzy model to detect if a true fall has occurred. The geometrical characteristics of the ROI corresponding to the detected person, in addition to the velocity of its change, serve as fall indicators. Additionally, the fuzzy model enables to avoid certain limitations in parameter evaluation, and to make smoother and more flexible decisions, on the other hand.

The proposed elderly-oriented fall detection and inactivity monitoring system has been tested for a number of situations, including the identification of true and false falls. The system is also capable of claiming for medical assistance when the fall persists in time.

Acknowledgements. This work was partially supported by Spanish Ministerio de Ciencia e Innovación / FEDER under project TIN2010-20845-C03-01, by Spanish Ministerio de Industria, Turismo y Comercio / FEDER (Plan Avanza 2) under project TSI-020100-2010-261, and by Spanish Junta de Comunidades de Castilla-La Mancha / FEDER under project PII2I09-0069-0994.

References

1. Anderson, D., Keller, J.M., Skubic, M., Chen, X., He, Z.: Recognizing falls from silhouettes. In: Proceedings of the IEEE Engineering in Medicine and Biology Society Conference, pp. 6388–6391 (2006)
2. Castillo, J.C., Serrano-Cuerda, J., Fernández-Caballero, A.: Robust People Segmentation by Static Infrared Surveillance Camera. In: García-Pedrajas, N., Herrera, F., Fyfe, C., Benítez, J.M., Ali, M. (eds.) IEA/AIE 2010. LNCS, vol. 6096, pp. 348–357. Springer, Heidelberg (2010)
3. Chang, S.-L., Yang, F.-T., Wu, W.-P., Cho, Y.-A., Chen, S.-W.: Nighttime pedestrian detection using thermal imaging based on HOG feature. In: Sokolova, M.V., Fernández-Caballero, A. (eds.) International Conference on System Science and Engineering, pp. 694–698 (2011)

4. Davis, J.W., Sharma, V.: Robust detection of people in thermal imagery. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 4, pp. 713–716 (2004)
5. Doukas, C., Maglogiannis, I.: Emergency fall incidents detection in assisted living environments utilizing motion, sound, and visual perceptual components. IEEE Transactions on Information Technology in Biomedicine 15(2), 277–289 (2011)
6. Escobet, A., Nebot, A., Cellier, F.E.: Fault diagnosis system based on fuzzy logic: Application to a valve actuator benchmark. Journal of Intelligent and Fuzzy Systems 22(4), 155–171 (2011)
7. Fang, Y., Yamada, K., Ninomiya, Y., Horn, B.K.P., Masaki, I.: A shape-independent method for pedestrian detection with far-infrared images. IEEE Transactions on Vehicular Technology 53(6), 1679–1697 (2004)
8. Fernández-Caballero, A., Castillo, J.C., Serrano-Cuerda, J., Maldonado-Bascón, S.: Real-time human segmentation in infrared videos. Expert Systems with Applications 38(3), 2577–2584 (2011)
9. Fernández-Caballero, A., Castillo, J.C., Martínez-Cantos, J., Martínez-Tomás, R.: Optical flow or image subtraction in human detection from infrared camera on mobile robot. Robotics and Autonomous Systems 58(12), 1273–1281 (2010)
10. Khawandi, S., Daya, B., Chauvet, P.: Implementation of a monitoring system for fall detection in elderly healthcare. Procedia Computer Science 3, 216–220 (2011)
11. Li, J., Gong, W., Li, W., Liu, X.: Robust pedestrian detection in thermal infrared imagery using the wavelet transform. Infrared Physics and Technology 53(4), 267–273 (2010)
12. Liu, C.-L., Lee, C.-H., Lin, P.-M.: A fall detection system using k-nearest neighbor classifier. Expert Systems with Applications 37(10), 7174–7181 (2010)
13. Loomis, A.: Figure Drawing for All it's Worth, Viking Adult (1943)
14. Mo, H.-C., Leou, J.-J., Lin, C.-S.: Human behavior analysis using multiple 2d features and multicategory support vector machine. In: Proceedings of the IAPR Conference on Machine Vision Applications, pp. 46–49 (2009)
15. Moon, H., Chellappa, R., Rosenfeld, A.: Tracking of human activities using shapeencoded particle propagation. In: Processing of the International Conference on Image Processing, vol. 1, pp. 357–360 (2001)
16. Moreno-Garcia, J., Rodriguez-Benitez, L., Fernández-Caballero, A., López, M.T.: Video sequence motion tracking by fuzzification techniques. Applied Soft Computing 10(1), 318–331 (2010)
17. Otsu, N.: A threshold selection method from gray level histograms. IEEE Transactions on Systems, Man and Cybernetics 9(1), 62–66 (1979)
18. Rougier, C., Meunier, J.: Demo: fall detection using 3d head trajectory extracted from a single camera video sequence. In: Proceedings of the First International Workshop on Video Processing for Security (2006)
19. Schmidt, J., Fritsch, J., Kwolek, B.: Kernel particle filter for real-time 3d body tracking in monocular color images. In: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, pp. 567–572 (2006)
20. Spark! Fuzzy Logic Editor (2011), http://www.louderthanabomb.com/spark.htm
21. Tan, S.C., Lim, C.P.: Fuzzy ARTMAP and hybrid evolutionary programming for pattern classification. Journal of Intelligent and Fuzzy Systems 22(2), 57–68 (2011)
22. Wei, X., Chai, J.: Modeling 3d human poses from uncalibrated monocular images. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1873–1880 (2009)
23. Yu, M., Rhuma, A., Naqvi, S.M., Chambers, J.: Fall detection for the elderly in a smart room by using an enhanced one class support vector machine. In: Proceedings of the 17th International Conference on Digital Signal Processing, pp. 1–6 (2011)

Interactive Approach for Random Fuzzy Two-Level Programming Problems through Fractile Criterion Optimization under Possibility

Shimpei Matsumoto, Kosuke Kato, and Hideki Katagiri

Abstract. In this paper, we focus on two-level linear programming problems whose coefficients in objective functions are expressed as random fuzzy variables. An interactive decision making approach is proposed in order to obtain a good solution from the viewpoint of possibility and probability. Transforming the original random fuzzy two-level programming problem into a deterministic one following the proposed decision making model, a compromise solution for the decision maker at the upper level is derived through interactions in consideration of the cooperative relationship between decision makers.

1 Introduction

In the real world, we often encounter situations where there are two or more decision makers in an organization with a hierarchical structure, and they make decisions in turn or at the same time so as to optimize their objective functions. In particular, consider a case where there are two decision makers; one of the decision makers first makes a decision, and then the other who knows the decision of the opponent makes a decision. Such a situation is formulated as a two-level programming problem. In the context of two-level programming, the decision maker at the upper level first specifies a strategy, and then the decision maker at the lower level specifies a strategy so as to optimize the objective with full knowledge of the action of the decision maker at the upper level. In conventional multi-level mathematical programming

Shimpei Matsumoto · Kosuke Kato
Hiroshima Institute of Technology, 2-1-1 Miyake, Saeki-ku, Hiroshima, 731-5193 Japan
e-mail: s.matsumoto.gk@it-hiroshima.ac.jp, k.katoh.me@it-hiroshima.ac.jp

Hideki Katagiri
Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527 Japan
e-mail: katagiri-h@hiroshima-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 225–234.
springerlink.com

models employing the solution concept of Stackelberg equilibrium, it is assumed that there is no communication among decision makers, or they do not make any binding agreement even if there exists such communication [4, 30]. Compared with this, for decision making problems in such as decentralized large firms with divisional independence, it is quite natural to suppose that there exists communication and some cooperative relationship among the decision makers.

Lai [21] and Shih et al. [29] proposed solution concepts for two-level linear programming problems or multi-level ones such that decisions of decision makers in all levels are sequential and all of the decision makers essentially cooperate with each other. Sakawa et al. have proposed interactive fuzzy programming for two-level or multi-level linear programming problems to obtain a satisfactory solution for decision makers [28]. The subsequent works on two-level or multi-level programming have been appearing [25, 3, 1, 26, 27].

However, to utilize two-level programming for resolution of conflict in decision making problems in real-world decentralized organizations, it is important to realize that simultaneous considerations of both fuzziness and randomness would be required. Fuzzy random variables, first introduced by Kwakernaak [20], have been developing, and an overview of the developments of fuzzy random variables was found in [6]. A brief survey of major fuzzy stochastic programming models was found in the paper by Luhandjula [24]. As we look at recent developments in the fields of fuzzy stochastic or fuzzy random programming, we can see continuing advances [16, 15, 2, 18, 19, 17]. It should be noted here that a fuzzy random variable is used to deal with the ambiguity of realized values of random variables. On the other hand, the concept of random fuzzy variables, focusing on the ambiguity of parameters characterizing random variables, has been proposed by Liu [23]. Intuitively speaking, a random fuzzy variable is an extended mathematical concept of random variable in the sense that it is defined as a fuzzy set defined on a universal set of random variables. For instance, the random variables with fuzzy mean values are represented with random fuzzy variables. Random fuzzy variables draw attention as a new tool for decision making under random fuzzy environments [12, 23, 8, 7, 31]. In particular, Katagiri et al. proposed an interactive fuzzy programming approach for random fuzzy two-level programming problems [13]. Although the decision makers need to determine membership functions quantifying their own fuzzy goals in their approach [13], in practical use, it would be difficult for the decision maker to easily determine the membership function which sufficiently reflect its preference.

Under these circumstances, in this paper, assuming cooperative behavior of the decision makers, we propose interactive approach for decision making problems in hierarchical organizations under random and fuzzy environments. To deal with the formulated two-level linear programming problems involving random fuzzy variables, a possibility-stochastic two-level linear programming problem is defined. The possibility-stochastic two-level linear programming problem is transformed into a deterministic one following the fractile criterion optimization model [11]. Then, an interactive algorithm is constructed to obtain a compromise solution for the decision maker at the upper level in consideration of the cooperative relationship between decision makers is presented.

2 Random Fuzzy Two-Level Linear Programming Problems

A random fuzzy variable, first introduced by Liu [23], is defined as:

Definition 1 (Random fuzzy variable [23]). A random fuzzy variable is a function ξ from a possibility space $(\Theta, P(\Theta), Pos)$ to a collection of random variables. An n-dimensional random fuzzy vector $n = (\xi_1, \xi_2, \ldots, \xi_n)$ is an n-tuple of random fuzzy variables $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$.

Intuitively speaking, a random fuzzy variable is an extended mathematical concept of random variable in the sense that it is defined as a fuzzy set defined on a universal set of random variables. For instance, the random variables with fuzzy mean values are represented with random fuzzy variables. It should be noted here that a random fuzzy variable is different from a fuzzy random variable [6, 15, 22, 24], which is used to deal with the ambiguity of realized values of random variables, not focusing on the ambiguity of parameters characterizing random variables like random fuzzy variables.

In this paper, as in [13], we consider the following two-level linear programming problem involving random fuzzy variable coefficients in objective functions formulated as:

$$\left.\begin{array}{ll} \underset{\text{for DM1}}{\text{minimize}} & z_1(\mathbf{x}_1, \mathbf{x}_2) = \bar{\tilde{\mathbf{C}}}_{11}\mathbf{x}_1 + \bar{\tilde{\mathbf{C}}}_{12}\mathbf{x}_2 \\ \underset{\text{for DM2}}{\text{minimize}} & z_2(\mathbf{x}_1, \mathbf{x}_2) = \bar{\tilde{\mathbf{C}}}_{21}\mathbf{x}_1 + \bar{\tilde{\mathbf{C}}}_{22}\mathbf{x}_2 \\ \text{subject to} & A_1\mathbf{x}_1 + A_2\mathbf{x}_2 \leq \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0} \end{array}\right\}. \tag{1}$$

It should be emphasized here that randomness and fuzziness of the coefficients are denoted by "dash above" and "wave above" i.e., "$\bar{\ }$" and "$\tilde{\ }$", respectively. In this formulation, $\mathbf{x}_1$ is an n_1 dimensional decision variable column vector for the DM at the upper level (DM1), $\mathbf{x}_2$ is an n_2 dimensional decision variable column vector for the DM at the lower level (DM2), $z_1(\mathbf{x}_1, \mathbf{x}_2)$ is the objective function for DM1 and $z_2(\mathbf{x}_1, \mathbf{x}_2)$ is the objective function for DM2. In (1), $\bar{\tilde{\mathbf{C}}}_{lj}$, $l = 1, 2$, $j = 1, 2$ are vectors whose elements $\bar{\tilde{C}}_{ljk}$, $k = 1, 2, \ldots, n_j$ are random fuzzy variables which are normal random variables with ambiguous mean values. In particular, as in [12, 13], we assume that the probability density function of $\bar{\tilde{C}}_{ljk}$ is formally expressed as

$$f_{ljk}(z) = \frac{1}{\sqrt{2\pi}\sigma_{ljk}} \exp\left(-\frac{(z - \tilde{M}_{ljk})^2}{2\sigma_{ljk}^2}\right), \tag{2}$$

where $\tilde{M}_{ljk}$ is an L-R fuzzy number characterized by the following membership function (see Figure 1):

$$\mu_{\tilde{M}_{ljk}}(\tau) = \begin{cases} L\left(\dfrac{m_{ljk} - \tau}{\lambda_{ljk}}\right) & (m_{ljk} \geq \tau) \\ R\left(\dfrac{\tau - m_{ljk}}{\rho_{ljk}}\right) & (m_{ljk} < \tau). \end{cases} \tag{3}$$

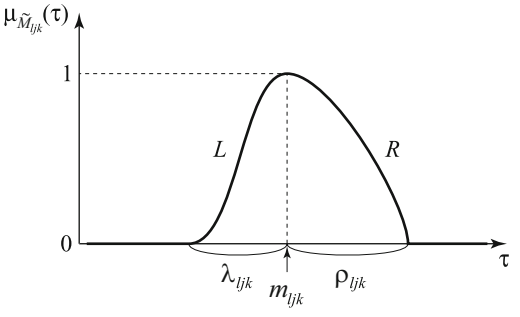


Fig. 1 Membership function $\mu_{\tilde{M}_{ljk}}(\cdot)$.

Functions L and R are called reference functions or shape functions which are non-increasing upper semi-continuous functions $[0,\infty) \to [0,1]$.

Random fuzzy two-level linear programming problems formulated as (1) are often seen in actual decision making situations. For example, consider a supply chain planning where the distribution center (DM1) and the production part (DM2) hope to minimize the distribution cost and the production cost, respectively. Since coefficients of these objective functions are often affected by the economic conditions varying at random, they can be regarded as random variables. In addition, since the statistical parameters of them are often ambiguous and estimated by fuzzy numbers, they are expressed by random fuzzy variables.

In the following, we transform (1) into a deterministic form in the similar manner in [12, 13].

When $\tilde{\bar{C}}_{ljk}$ is a random fuzzy variable characterized by (2) and (3), the membership function of $\tilde{\bar{C}}_j$ is expressed as

$$\mu_{\tilde{\bar{C}}_{ljk}}(\bar{\gamma}_{ljk}) = \sup_{s_{ljk}} \{\mu_{\tilde{M}_{ljk}}(s_{ljk}) | \bar{\gamma}_{ljk} \sim N(s_{ljk}, \sigma^2_{ljk})\}, \tag{4}$$

where $\bar{\gamma}_{ljk} \in \Gamma$, and Γ is a universal set of normal random variables. Each membership function value $\mu_{\tilde{\bar{C}}_{ljk}}(\bar{\gamma}_j)$ is interpreted as a degree of possibility or compatibility that $\tilde{\bar{C}}_{ljk}$ is equal to $\bar{\gamma}_{ljk}$.

Then, applying the results shown by Liu [23], the objective function $\tilde{\bar{\mathbf{C}}}_l\mathbf{x}$ is defined as a random fuzzy variable characterized by the following probability density function:

$$f_{\tilde{\bar{\mathbf{C}}}_l\mathbf{x}}(z) = \frac{1}{\sqrt{2\pi\mathbf{x}^T V_l \mathbf{x}}} \exp\left(-\frac{\left(z - \tilde{\mathbf{M}}_l\mathbf{x}\right)^2}{2\mathbf{x}^T V_l \mathbf{x}}\right), \tag{5}$$

where $\tilde{\mathbf{M}}_l = (\tilde{M}_1, \tilde{M}_2, \ldots, \tilde{M}_{n_1+n_2})$ and V_l is the covariance matrix for $\tilde{\bar{C}}_{ljk}$, $j = 1,2$, $k = 1,2,\ldots,n_j$. To be more specific, the membership function of $\tilde{\mathbf{M}}_l\mathbf{x}$ is expressed as:

$$\mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\tau) = \begin{cases} L\left(\dfrac{\mathbf{m}_l\mathbf{x}-\tau}{\lambda_l\mathbf{x}}\right) & (\mathbf{m}_l\mathbf{x} \geq \tau) \\ R\left(\dfrac{\tau-\mathbf{m}_l\mathbf{x}}{\rho_l\mathbf{x}}\right) & (\mathbf{m}_l\mathbf{x} < \tau) \end{cases} \tag{6}$$

where $\mathbf{m}_l = (m_1, m_2, \ldots, m_{n_1+n_2})$, $\lambda_l = (\lambda_1, \lambda_2, \ldots, \lambda_{n_1+n_2})$, $\rho_l = (\rho_1, \rho_2, \ldots, \rho_{n_1+n_2})$.

As an extension of (4) based on (5) and (6), we obtain

$$\mu_{\tilde{\bar{\mathbf{C}}}_l\mathbf{x}}(\bar{u}_l) = \sup_{\mathbf{s}_l}\left\{\mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \,\middle|\, \bar{u}_l \sim N\left(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l\mathbf{x}\right)\right\} \tag{7}$$

where $\mathbf{s}_l = (s_{l11}, \ldots, s_{l1n_1}, s_{l21}, \ldots, s_{l2n_2})$.

Considering that the mean values of random variables are often ambiguous and estimated as fuzzy numbers using experts' knowledge based on their experiences, the coefficients are expressed by random fuzzy variables. Then, such a supply chain planning problem can be formulated as a two-level linear programming problem involving random fuzzy variable coefficients like (1).

In stochastic programming, basic optimization criterion is to simply optimize the expectation of objective function values or to decrease their fluctuation as little as possible from the viewpoint of stability. In contrast to these types of optimizing approaches, the fractile model or Kataoka's model [11] has been proposed when the decision maker wishes to optimize a permissible level under the guaranteed probability that the objective function value is better than or equal to the permissible level.

On the other hand, in fuzzy programming, possibilistic programming [9] is one of methodologies for decision making under existence of ambiguity of the coefficients in objective functions and constraints.

In this research, we consider a random fuzzy two-level programming model, similar to [12, 13], in order to simultaneously optimize both possibility and probability with the attained objective function values of DMs, by extending both viewpoints of stochastic programming and possibilistic programming.

To begin with, let us express the probability $\Pr\left(\omega \,\middle|\, \tilde{\bar{\mathbf{C}}}_l(\omega)\mathbf{x} \leq f_l\right)$ as a fuzzy set $\tilde{P}_l$ and define the membership function of $\tilde{P}_l$ as follows:

$$\mu_{\tilde{P}_l}(p_l) = \sup_{\bar{u}_l}\left\{\mu_{\tilde{\bar{\mathbf{C}}}_l\mathbf{x}}(\bar{u}_l) \,\middle|\, p_l = \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l)\right\}. \tag{8}$$

From (7) and (8), we obtain

$$\mu_{\tilde{P}_l}(p_l) = \sup_{\mathbf{s}_l}\left\{\mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \,\middle|\, p_l = \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l), \bar{u}_l \sim N\left(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l\mathbf{x}\right)\right\}.$$

Assuming that each of DMl, $l = 1, 2$ has a goal G_l for the attained probability expressed as "$\tilde{P}_l$ should be greater than or equal to some value $\hat{p}_l(\geq 0.5)$." Then, the possibility that $\tilde{P}_l$ satisfies the goal G_l is expressed as

$$\Pi_{\tilde{P}_l}(G_l) \stackrel{\triangle}{=} \sup_{p_l}\{\mu_{\tilde{P}_l}(p_l) \mid p_l \geq \hat{p}_l\}.$$

Then, as a new optimization model in two-level programming problems with random fuzzy variables, we consider the following possibility-based fractile criterion optimization model:

$$\left.\begin{array}{ll} \underset{\text{for DM1}}{\text{minimize}} & f_1 \\ \underset{\text{for DM2}}{\text{minimize}} & f_2 \\ \text{subject to} & \Pi_{\tilde{P}_l}(G_l) \geq h_l,\ l = 1,2 \\ & A_1\mathbf{x}_1 + A_2\mathbf{x}_2 \leq \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0} \end{array}\right\} \tag{9}$$

where h_1 and h_2 are constants.

Then we can derive the following theorem:

Theorem 1. *Let Φ denote the distribution function of $N(0,1)$ and $K_{\hat{p}_l} \stackrel{\triangle}{=} \Phi^{-1}(\hat{p}_l)$. Then, $\Pi_{\tilde{P}_l}(G_l) \geq h_l$ in problem (9) is equivalently transformed into the following condition*

$$(\mathbf{m}_l - L^*(h_l)\lambda_l)\mathbf{x} + K_{\hat{p}_l}\mathbf{x}^T V_l \mathbf{x} \leq f_l$$

In this condition, $L^*(h_l)$ is a pseudo inverse function defined as $L^*(h_l) = \sup\{t \mid L(t) \geq h_l\}$.

Proof. The constraint $\Pi_{\tilde{P}_l}(G_l) \geq h_l,\ l = 1,2$ in problem (9) is equivalently transformed as follows:

$$\begin{aligned}
&\Pi_{\tilde{P}_l}(G_l) \geq h_l \\
&\Leftrightarrow \exists p_l :\ \mu_{\tilde{P}_l}(p_l) \geq h_l,\ p_l \geq \hat{p}_l \\
&\Leftrightarrow \exists p_l :\ \sup_{\mathbf{s}_l} \mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \geq h_l, p_l \geq \hat{p}_l,\ p_l = \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l), \bar{u}_l \sim N\left(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}\right) \\
&\Leftrightarrow \exists \mathbf{s}_l :\ \mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \geq h_l, \bar{u}_l \sim N\left(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}\right), \Pr(\omega | \bar{u}_l(\omega) \leq f_l) \geq \hat{p}_l
\end{aligned}$$

From the definition of $\mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x})$, the set of $\mathbf{s}_l\mathbf{x}$ satisfying $\mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \geq h_l$ is an interval $[\tau_l^L, \tau_l^R]$ where τ_l^L and τ_l^R fulfill $L\left(\dfrac{\mathbf{m}_l\mathbf{x} - \tau_l^L}{\lambda_l\mathbf{x}}\right) = h_l$ and $R\left(\dfrac{\tau_l^R - \mathbf{m}_l\mathbf{x}}{\rho_l\mathbf{x}}\right) = h_l$, respectively. Therefore,

$$\begin{aligned}
&\exists \mathbf{s}_l :\ \mu_{\tilde{\mathbf{M}}_l\mathbf{x}}(\mathbf{s}_l\mathbf{x}) \geq h_l, \bar{u}_l \sim N\left(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}\right), \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l) \geq \hat{p}_l \\
&\Leftrightarrow \exists \mathbf{s}_l :\ \mathbf{s}_l\mathbf{x} \in [(\mathbf{m}_l - L^*(h_l)\lambda_l)\mathbf{x},\ (\mathbf{m}_l + R^*(h_l)\rho_l)\mathbf{x}], \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l) \geq \hat{p}_l, \\
&\quad \bar{u}_l \sim N(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}) \\
&\Leftrightarrow \exists \mathbf{s}_l :\ \Pr\left(\omega \mid \bar{u}_l^R(\omega) \leq f_l\right) \leq \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l) \leq \Pr\left(\omega \mid \bar{u}_l^L(\omega) \leq f_l\right), \\
&\quad \Pr(\omega \mid \bar{u}_l(\omega) \leq f_l) \geq \hat{p}_l, \bar{u}_l \sim N(\mathbf{s}_l\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}), \\
&\quad \bar{u}_l^L \sim N((\mathbf{m}_l - L^*(h_l)\lambda_l)\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}), \bar{u}_l^R \sim N((\mathbf{m}_l + R^*(h_l)\rho_l)\mathbf{x}, \mathbf{x}^T V_l \mathbf{x}) \\
&\Leftrightarrow \Pr\left(\omega \mid \bar{u}_l^L(\omega) \leq f_l\right) \geq \hat{p}_l, \bar{u}_l^L \sim N((\mathbf{m}_l - L^*(h_l)\lambda_l)\mathbf{x}, \mathbf{x}^T V_l \mathbf{x})
\end{aligned}$$

$$\Leftrightarrow \Phi\left(\frac{f_l-(\mathbf{m}_l-L^*(h_l)\lambda_l)\mathbf{x}}{\sqrt{\mathbf{x}^T V_l \mathbf{x}}}\right) \geq \hat{p}_l$$
$$\Leftrightarrow (\mathbf{m}_l-L^*(h_l)\lambda_l)\mathbf{x}+K_{\hat{p}_l}\sqrt{\mathbf{x}^T V_l \mathbf{x}} \leq f_l$$

From Theorem 1, problem (9) is transformed into the following problem:

$$\left.\begin{array}{ll} \underset{\text{for DM1}}{\text{minimize}} & f_1 \\ \underset{\text{for DM2}}{\text{minimize}} & f_2 \\ \text{subject to} & (\mathbf{m}_l-L^*(h_l)\lambda_l)\mathbf{x}+K_{\hat{p}_l}\sqrt{\mathbf{x}^T V_l \mathbf{x}} \leq f_l,\ l=1,2 \\ & A_1\mathbf{x}_1+A_2\mathbf{x}_2 \leq \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0} \end{array}\right\}. \tag{10}$$

Finally, (10) is reduced to the following problem:

$$\left.\begin{array}{ll} \underset{\text{for DM1}}{\text{minimize}} & (\mathbf{m}_1-L^*(h_1)\lambda_1)\mathbf{x}+K_{\hat{p}_1}\sqrt{\mathbf{x}^T V_1 \mathbf{x}} \\ \underset{\text{for DM2}}{\text{minimize}} & (\mathbf{m}_2-L^*(h_2)\lambda_2)\mathbf{x}+K_{\hat{p}_2}\sqrt{\mathbf{x}^T V_2 \mathbf{x}} \\ \text{subject to} & A_1\mathbf{x}_1+A_2\mathbf{x}_2 \leq \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0} \end{array}\right\}, \tag{11}$$

where $K_{\hat{p}_l} \geq 0$ because $\hat{p}_l \geq 1/2$. For simplicity, we express

$$Z_l^F(\mathbf{x}_1,\mathbf{x}_2) = (\mathbf{m}_l-L^*(h_l)\lambda_l)\mathbf{x}+K_{\hat{p}_l}\sqrt{\mathbf{x}^T V_l \mathbf{x}}. \tag{12}$$

Now we construct the following interactive algorithm to derive a compromise solution for the decision maker at the upper level in consideration of the cooperative relationships between DM1 and DM2 based on an improved satisficing trade-off method [10].

Interactive Algorithm to Derive a Cooperative Compromise Solution

Step 1. In order to calculate the individual minimum of f_1 and f_2, solve the following problems:

$$\left.\begin{array}{ll} \text{minimize} & Z_l^F(\mathbf{x}_1,\mathbf{x}_2) \\ \text{subject to} & A_1\mathbf{x}_1+A_2\mathbf{x}_2 \leq \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0} \end{array}\right\},\ l=1,2 \tag{13}$$

Let $\mathbf{x}_{l,\min}$ and $Z_{l,\min}^F$ be the optimal solution to (13) and the minimal objective function value to (13), respectively. Observing that (13) are convex programming problems, they can be easily solved by some convex programming technique like sequential quadratic programming methods [5].

Step 2. Ask DMs to specify probability aspiration levels $\hat{p}_l \in (0,1)$, $l=1,2$, possibility aspiration levels $h_l \in (0,1)$, $l=1,2$ and objective aspiration levels Z_l^{Fa}, $l=1,2$.

Step 3. The following problem is solved for obtaining a solution which minimizes the larger value of normalized differences between $Z_l^F(\mathbf{x})$ and Z_l^{Fa}, $l = 1,2$:

$$\left.\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \dfrac{Z_l^F(\mathbf{x}) - Z_l^{Fa}}{|Z_{l,\min}^F|} \le v,\ l = 1,2 \\ & A_1\mathbf{x}_1 + A_2\mathbf{x}_2 \le \mathbf{b} \\ & \mathbf{x}_1 \ge \mathbf{0},\ \mathbf{x}_2 \ge \mathbf{0} \end{array}\right\}. \tag{14}$$

Since this problem is also a nonlinear convex programming problem, we can solve it by some convex programming technique.

Step 4. DM1 is supplied with the current values of $Z_1^F(\mathbf{x}^*)$ and $Z_2^F(\mathbf{x}^*)$ for the optimal solution $\mathbf{x}^*$ calculated in step 3. If DM1 is satisfied with the current function values, the interaction process is terminated. If DM1 is not satisfied and desires to update $\hat{p}_l$, $l = 1,2$, h_l, $l = 1,2$ and/or Z_l^{Fa}, $l = 1,2$, ask DM1 to update h_ls and/or Z_l^{Fa}s and go to step 3. Otherwise, ask DM1 to specify $\beta_l \in [0,1]$, $l = 1,2$, the degree as a goal of the constraint $Z_l^F(\mathbf{x}) \le Z_l^{Fa}$. Considering the priority of DM1 over DM2, β_1 should be less than β_2. Then, (14) is extended as:

$$\left.\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \dfrac{Z_l^F(\mathbf{x}) - Z_l^{Fa}}{|Z_{l,\min}^F|} \le \beta_l v,\ l = 1,2 \\ & A_1\mathbf{x}_1 + A_2\mathbf{x}_2 \le \mathbf{b} \\ & \mathbf{x}_1 \ge \mathbf{0},\ \mathbf{x}_2 \ge \mathbf{0} \end{array}\right\}. \tag{15}$$

If $\beta_l = 1$, $Z_l^F(\mathbf{x}) \le Z_l^{Fa}$ is a goal completely. On the other hand, if $\beta_l = 0$, $Z_l^F(\mathbf{x}) \le Z_l^{Fa}$ is a constraint completely.

Step 5. For the specified values of $\hat{p}_l$, h_l, Z_l^{Fa} and β_l, $l = 1,2$, solve (15). Since (15) is a convex programming problem, they can be easily solved by some convex programming technique.

Step 6. DM1 is supplied with the current values of $Z_1^F(\mathbf{x}^*)$, $Z_2^F(\mathbf{x}^*)$. If DM1 is satisfied with the current objective function values for the optimal solution $\mathbf{x}^*$, the interaction process is terminated. Otherwise, ask DM1 to update the probability aspiration levels $\hat{p}_l$, $l = 1,2$, the possibility aspiration levels h_l, $l = 1,2$, objective aspiration levels Z_l^{Fa}, and/or goal degrees β_l, $l = 1,2$ and return to step 5.

It should be noted that the proposed interactive algorithm is different from that in [13] with respect to the following points: 1) the decision makers do not have to determine membership functions quantifying their own fuzzy goals (in practical use, it would be difficult for the decision maker to easily determine the membership function which sufficiently reflect its preference), 2) the inequality $Z_l^F(\mathbf{x}) \le Z_l^{Fa}$ where Z_l^{Fa} is an objective aspiration level can be handled as a constraint or a goal depending on a parameter value β_l, 3) we can carry out similar interactions to [13] by setting $\beta_1 = 0$ and $\beta_2 = 1$.

3 Conclusion

In this paper, assuming cooperative behavior of the decision makers, interactive decision making methods in hierarchical organizations under random fuzzy environments have been considered. For the formulated random fuzzy two-level linear programming problems, through the introduction of the possibility-based fractile model as a new decision making model, we has shown that the original random fuzzy two-level programming problem is reduced to a deterministic one. In order to obtain a cooperative compromise solution for the decision maker at the upper level in consideration of the cooperative relationship between decision makers, we have constructed an interactive algorithm in which each of optimal solutions of all problems can be analytically obtained by some techniques for solving convex programming problems.

We will consider applications of the proposed method to real-world decision making problems in decentralized organizations together with extensions to other stochastic programming models in the near future. Furthermore, we will extend the proposed models and concepts to noncooperative random fuzzy two-level linear programming problems.

References

1. Abo-Sinna, M.A., Baky, I.A.: Interactive balance space approach for solving multi-level multi-objective programming problems. Inform. Sciences 177, 3397–3410 (2007)
2. Ammar, E.E.: On solutions of fuzzy random multiobjective quadratic programming with applications in portfolio problem. Inform. Sciences 178, 468–484 (2008)
3. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization 52, 333–359 (2003)
4. Edmunds, T.A., Bard, J.F.: Algorithms for nonlinear bilevel mathematical programs. IEEE T. Syst. Man Cyb. 21, 83–89 (1991)
5. Fletcher, R.: Practical Methods of Optimization, 2nd edn. Willey, New York (2000)
6. Gil, M.A., Lopez-Diaz, M., Ralescu, D.A.: Overview on the development of fuzzy random variables. Fuzzy Set. Syst. 157, 2546–2557 (2006)
7. Hasuike, T., Katagiri, H., Ishii, H.: Portfolio selection problems with random fuzzy variable returns. Fuzzy Set. Syst. 160, 2579–2596 (2009)
8. Huang, X.: A new perspective for optimal portfolio selection with random fuzzy returns. Inform. Sciences 177, 5404–5414 (2007)
9. Inuiguchi, M., Ramik, J.: Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Set. Syst. 111, 3–28 (2000)
10. Kamenoi, T., Sugimoto, H., Nakayama, H.: A study on improved satisficing trade-off method for structural design. J. Jpn. Soc. Civ. Eng. 441, 117–126 (1992)
11. Kataoka, S.: A stochastic programming model. Econometorica 31, 181–196 (1963)
12. Katagiri, H., Ishii, H., Sakawa, M.: Linear programming problems with random fuzzy variable coefficients. In: Proceedings of 5th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty, vol. 1, pp. 55–58 (2002)

13. Katagiri, H., Niwa, K., Kubo, D., Hasuike, T.: Interactive random fuzzy two-level programming through possibility-based fractile criterion optimality. In: Proceedings of The International Multi Conference of Engineers and Computer Scientists 2010, vol. 3, pp. 2113–2118 (2010)
14. Katagiri, H., Sakawa, M., Kato, K., Nishizaki, I.: A fuzzy random multiobjective 0-1 programming based on the expectation optimization model using possibility and necessity measures. Math. Comput. Model. 40, 411–421 (2004)
15. Katagiri, H., Sakawa, M., Kato, K., Nishizaki, I.: Interactive multiobjective fuzzy random linear programming: maximization of possibility and probability. Eur. J. Oper. Res. 188, 530–539 (2008)
16. Kato, K., Katagiri, H., Sakawa, M., Wang, J.: Interactive fuzzy programming based on a probability maximization model for two-level stochastic linear programming problems. Electron. Comm. Jpn. 89(3), 33–42 (2006)
17. Kato, K., Sakawa, M.: An interactive fuzzy satisficing method based on variance minimization under expectation constraints for multiobjective stochastic linear programming problems. Soft Comput. 15, 131–138 (2011)
18. Kato, K., Sakawa, M., Katagiri, H.: Interactive fuzzy programming for two-level stochastic linear programming problems through expectation and variance models. KES Journal 13, 111–118 (2009)
19. Kato, K., Sakawa, M., Katagiri, H., Perkgoz, C.: An interactive fuzzy satisficing method based on fractile criterion optimization for multiobjective stochastic integer programming problems. Expert Syst. Appl. 37, 6012–6017 (2010)
20. Kwakernaak, H.: Fuzzy random variables - I. definitions and theorems. Inform. Sciences 15, 1–29 (1978)
21. Lai, Y.J.: Hierarchical optimization: a satisfactory solution. Fuzzy Set. Syst. 77, 321–325 (1996)
22. Liu, B.: Fuzzy random chance-constrained programming. IEEE T. Fuzzy Syst. 9, 713–720 (2001)
23. Liu, B.: Theory and Practice of Uncertain Programming. Physica Verlag, Heidelberg (2002)
24. Luhandjula, M.K.: Fuzzy stochastic linear programming: survey and future research directions. Eur. J. Oper. Res. 174, 1353–1367 (2006)
25. Patriksson, M., Wynter, L.: Stochastic mathematical programs with equilibrium constraints. OR Letters 25, 159–167 (1999)
26. Pramanik, S., Roy, T.K.: Fuzzy goal programming approach to multilevel programming problems. Eur. J. Oper. Res. 176, 1151–1166 (2007)
27. Roghanian, E., Sadjadi, S.J., Aryanezhad, M.B.: A probabilistic bi-level linear multiobjective programming problem to supply chain planning. Appl. Math. Comput. 188, 786–800 (2007)
28. Sakawa, M., Nishizaki, I.: Cooperative and Noncooperative Multi-Level Programming. Springer, New York (2009)
29. Shih, H.S., Lai, Y.J., Lee, E.S.: Fuzzy approach for multi-level programming problems. Comput. Oper. Res. 23, 73–91 (1996)
30. Shimizu, L., Ishizuka, Y., Bard, J.F.: Nondifferentiable and Two-Level Mathematical Programming. Kluwer Academic Publishers, Boston (1997)
31. Xu, J., Liu, Q., Wang, R.: A class of multi-objective supply chain networks optimal model under random fuzzy environment and its application to the industry of Chinese liquor. Inform. Sciences 178, 2022–2043 (2008)

Main Branch Decision Tree Algorithm for Yield Enhancement with Class Imbalance

Chia-Yu Hsu*, Chen-Fu Chien, and Yai-Chun Lai

Abstract. Continuous yield enhancement is crucial for semiconductor companies that are driven by Moore's Law for technology migration. This study aims to develop a manufacturing intelligence approach for trouble shooting and low yield correlation with imbalanced classes. In particular, Engineering Data Analysis (EDA) system is an off-line analysis system that can be used for trouble-shooting and yield enhancement. While making trouble shooting in semiconductor industry, engineers can not only use the expert knowledge on physics or electronics to answer the problem because of numerous relevant analysis factors. However, little research has been done on analyzing class imbalance problem to extract useful rules for yield enhancement. This study proposed a main branch decision tree (MBDT) algorithm that modifies the criteria of tree growth rather than using accuracy-based methods. An empirical study was conducted in a semiconductor company to evaluate validity. The results provide references for engineers to quickly identify the assignable causes and diagnose the problem of low yield.

Keywords: manufacturing intelligence, yield enhancement, main branch decision tree algorithm, class imbalance, semiconductor manufacturing.

1 Introduction

In order to keep the competitive advantages in semiconductor industry, semiconductor fabrication needs effective analytical method to deal with relevant data

Chia-Yu Hsu
Department of Information Management, Yuan Ze University, 32003, Chungli, Taiwan
e-mail: cyhsu@saturn.yzu.edu.tw

Chen-Fu Chien · Yai-Chun Lai
Department of Industrial Engineering and Engineering Management
National Tsing Hua University, 30013, Hsinchu, Taiwan

* Corresponding author.

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 235–244.
springerlink.com

analysis and decision problem for production improvement, cost reduction and yield enhancement. In particular, quick response to process excursion is crucial to enhance yield in semiconductor industry. However, the yield is affected by many manufacturing factors that are often interrelated due to the complex manufacturing processes, and numerous equipments and parameters in semiconductor industry. It is difficult for engineers to rapidly identify possible root causes of low yield by own domain knowledge only.

With the advance in information technology, a large amount of process data will be automatically or semi-automatically recorded and stored in the engineering database for process monitoring, fault diagnosis and manufacturing management during the fabrication process. Manufacturing intelligence (MI) that can be extracted from huge amount of manufacturing-related database and data value development has become an effective approach to enhance the decision quality for competitive high-tech industries such as semiconductor manufacturing [6, 7, 12]. The MI can be employed to assist data value development, manufacturing knowledge discovery, and explanation to maintain the competitive advantages.

In general, the number of defective product tends to be smaller than the number of non-defective products. This type of data is called class imbalance data. The typical classification methods are unable to classify correctly unseen instances from the minority class without considering the class distribution [13]. Most of existing studies to improve classification result of class imbalance are used to change the class distribution [10]. However, little research has been done on analyzing class imbalance problem to extract useful rules for yield enhancement.

This study aims to develop a manufacturing intelligence approach to extract useful rules for identifying assignable causes for low yield. In particular, a novel main branch decision tree (MBDT) algorithm is developed to explore the focus class (FC) for low yield. In order to modify the tree pruning in class imbalance problem, Relative-Purity (RP) and Focus Class Significance (FCSIG) are proposed for pre-pruning in tree growth and Focus-Information-Gain (FIG) is proposed for post-pruning. To validate the proposed approach, we also conducted an empirical study from a leading semiconductor company. The extraction rules and patterns can provide the reference for engineers to shrink the range of possible causes and shorten the time of trouble shooting. The results demonstrated the practical visibility of proposed approach.

2 Fundamental

Continuous yield improvement is crucial for wafer fabrications to maintain competitive advantages as global competition continues to strengthen in semiconductor industry. The yield is defined as the ratio of the number of good units of a manufactured IC that meets all the performance, power, functionality, and quality specifications of the product to the total number of manufactured IC [3]. In particular, the high yield can obtain lower manufacturing cost and high profitability. Therefore, quickly detection for inherent yield loss and deal with the cause factors is essential for yield improvement. The yield loss on a wafer is due to the defect,

which can be grouped into random defects and systematic defects. The random defects are visual and typically caused by particles during wafer production. The systematic defects are typically caused by the variability of equipments, chambers, parameters in different processes [3].

Trouble shooting is to identify the root cause of low yield quickly for assurance of further wafer quality by analyzing off-line data including defect analysis, wafer acceptance test (WAT) and circuit probing (CP) test. The manufacturing process in semiconductor fabrication involves hundreds of processes and equipments and the defective wafers could be caused by front manufacturing process or particular equipment abnormality. Due to the complexity of wafer fabrication processes, it is difficult for engineers to rapidly find the potential root-causes based on domain knowledge and experience.

Owing to advances of information technologies and new applications, large amounts of data are recorded that can provide a rich resource for data mining and knowledge discovery from database. Decision makers may potentially use the information buried in the raw data to assist their decisions through data mining for possibly identifying the specific pattern of the data [1]. Data mining methods have been widely applied for yield improvement in high-tech industry [2, 5, 8, 9, 14].

The typical classification methods such like decision tree, Bayesian network and instance-based classifier are often biased towards the majority class and unable to classify correctly the unseen cases from the minority class [11, 13]. These classifiers attempt to achieve global optimization by reducing the error rate, rather than considering the data distribution into model construction. Evaluating the classification performance of imbalanced data is not done appropriately by standard accuracy/error rate measures. To improve the performance of classification method for class imbalance data, three methods including over-sampling, under-sampling, and learning by recognition are developed [10]. Over-sampling method is to increase the data from minority class and under-sampling method is to remove the data from the majority class. Both of over-sampling and under-sampling methods try to match the size of classes. The learning by recognition method ignores one of the two classes, altogether, by using a recognition-based instead of a discrimination-based inductive scheme.

3 Proposed Approach

3.1 Data Preparation

Amount of engineering data are automatically or semi-automatically recorded during manufacturing process. There are six types of data including WIP (wafer in process) data, metrology data, non-lot data, defect data, WAT test data, and CP test data. According to the collected data, Engineers can analyze these data to identify the relationship between the test measurements (metrology data, WAT data, and CP data) and WIP data. Because low yield problems are usually caused by specific equipment, chambers, or recipes in operation process and operation time, this study focuses on the CP data and WIP data for trouble shooting. In particular,

the bin items of CP test are specified as the response variable, and relative process stations involving operation machines and operation time are specified as predictors.

The collected data often includes noisy, missing and inconsistent data. Data preprocessing can improve the quality of the data and transform the data for appropriate format. This step includes data cleaning, data clustering and data treatment. Data cleaning is to remove the instances, outliers and the sample without CP yield and empty process data. Data clustering is to group the lots which have similar performance in yield and bin item of CP test by two-stage clustering method [4]. Furthermore, Kruskal-Wallis (K-W) test is applied to screen out significant attributes for group results of yield and target bin item of CP test. Then, the predictors which can be significant in both results are selected for the further MBDT model construction.

3.2 Main Branch Decision Tree

MBDT focuses on the class of low yield which is critical to extract meaningful rules associated with operation process information. MBDT is designed for categorical response variable and categorical predictors as shown in Fig. 3.1. The key differences between MBDT and other decision tree algorithms is that pre-pruning and post-pruning criteria are selected based on distribution of minority class. During the decision tree growing, the RP threshold and FCSIG threshold are used as stopping criterion (pre-pruning), and the FIG is used as post-pruning. We use the greedy splitting criterion for tree growth. In rules extraction step, the decision rules are derived for trouble shooting from the path of root node to each leaf node. If-Then rules show the relationships between low yield and operation process variables. The modeling results need to be interpreted and discussed with domain experts.

First, there have five parameters including merge-P-value, split-P-value, significance-P-value, RP threshold, and FCSIG threshold need to be determined by user. In particular, the merge-P-value is used to decide whether the significance between the category in a predictor reach the critical value. The split-P-value is used to decide the optimally subdivided groups of the predictor attribute. The significance-P-value is the threshold to check whether the significance of most appropriate split predictor reaches the critical value. If exceeds the critical value, predictor is selected and split the tree; otherwise, it will stop growing.

Moreover, the RP threshold is to measure the proportion of FC instances in the internal node between other classes. When value of RP is near 1, it means the node is less purity and need for advanced split to separate out the FC instances. If the value of RP near 0, it means that the node is more purity with other classes and can be set as lead node.

$$RP(C_f) = \frac{n_f}{n_{\bar{f}}},\ 0 \le RP(C_f) \le 1 \tag{3.1}$$

where $RP(C_f)$ represent the relative purity of FC (C_f) in each node. n_f is the number of instances which belong to the focus class in the node and $n_{\bar{f}}$ is the number of non-focus class in the node.

The FCSIG threshold is to determine whether the node can increase the information of FC in tree growth. FCSIG will change dynamically with the relationship between parent node and child node.

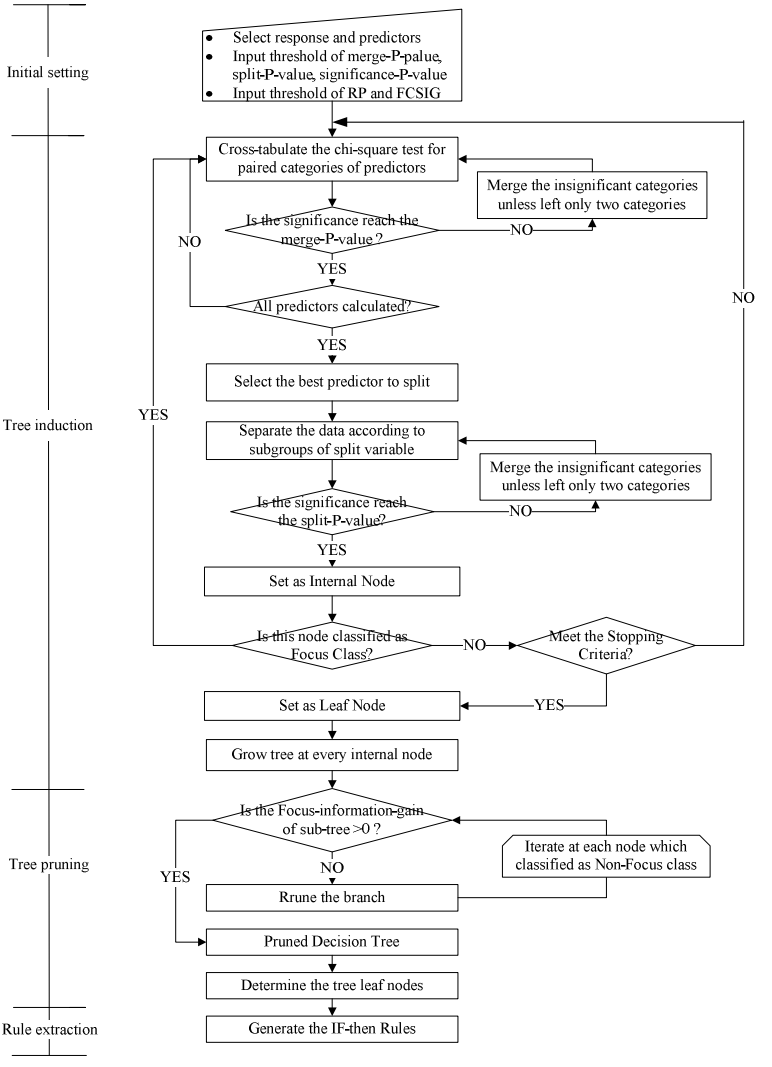


Fig. 3.1 Main branch decision tree (MBDT) algorithm

$$FCSIG(i) = \frac{n_f(\text{child})}{n_f(\text{parent})} \tag{3.2}$$

where $n_f(\text{child})$ is the number of samples which are belong to the FC in the child node and $n_f(\text{parent})$ is the number of samples which are belong to the FC in the parent node. Furthermore, FIG is developed to account the information from FC and to avoid the bias from accuracy-based pruning methods. The leaf is pruned while FIG is zero which means no useful information about FC in certain sub-tree.

$$FIG(t) = FC\_TP(t)/N(t) \tag{3.3}$$

where t is the sub-tree of internal node. FC_TP(t): the number of correct classified FC instances in the sub-tree. N(t): number of instances in the sub-tree.

4 Empirical Study

To validate the proposed approach, an empirical study was conducted from a semiconductor fabrication. The production products with low yield lots are collected. In particular, low CP yield problem are also highly biased bin 4 test data. Therefore, we try to find out the root causes equipments resulting in the abnormal condition of bin 4 item and yield.

4.1 Data Preparation

The possible root causes may be some machines at the process station or time periods, and total three months data in CP test data and WIP data were analyzed from the engineering database. The CP test data includes lots id, Target Bin test data of lots, yield data of lots, and the WIP data includes lot id, process station of lots, operation machine of lots, operation time-point of lots. There are 1260 lots are collected and there are 1315 relative process stations.

First, the CP test data and WIP data are integrated into a single dataset for following analysis. Next, the outlier and missing value in response are removed and the process stations which are not passed by any wafer are deleted. The remaining 774 samples and 800 process variables are used for low yield analysis. In particular, 90% instances (697 lots) are randomly selected as training data and the other 10% instances (77 lots) are selected as testing data.

Next, both CP yield and bin 4 value are standardized by mean and standard deviation and two stage clustering approach is used to cluster the 697 instances There are three clusters as shown in Fig. 4.1. In particular, Cluster#1 has 292 instances, Cluster#2 has 290 instances, and Cluster#3 has 115 instances. The basic statistics for each cluster are show in Fig. 4.2. There has high negative correlation between yield and bin 4 item for cluster#1 and cluster#3. We select the cluster#1 and cluster#3 as target classification variable.

K-W test is to find what attributes are significantly different between operation machines or operation machines combined with operation date according to yield and bin 4 item, respectively. The significance level is set as 0.05. If the p-value of an attribute in K-W test is less than 0.05, we can say that there is significant difference between levels in the attribute. Partial result of K-W test is listed in Table 4.1 and 99 significant predictors are selected for the MBDT construction as shown in Table 4.2.

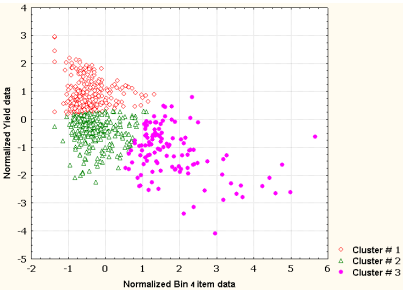


Fig. 4.1 Scatter plot of normalized CP yield data and bin 4 item data of each cluster

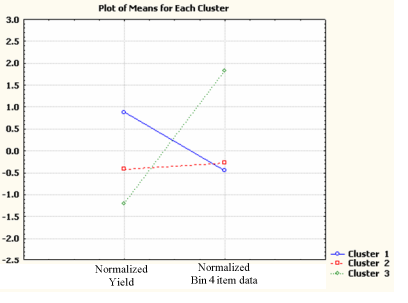


Fig. 4.2 Average of normalized CP yield data and bin 4 item data of each cluster

Table 4.1 Partial selected predictor after K-W test screening (99 predictors)

Var. No.	Predictor Name	p-value (for Yield)	p-value (for bin 4)
6	1B_DHL52_EQUIPMENT	0.0046	0.0022
18	1B_PH811_EQUIPMENT	0.0001	0.0007
34	1B_WHWC5_EQUIPMENT	0.0030	0.0046
54	1C_PSKX3	0.0135	0.0239
65	1D_D56L1_EQUIPMENT	0.0086	0.0076
67	1D_I5BH1_EQUIPMENT	0.0082	0.0061
70	1D_P5811	0.0388	0.0272
71	1D_P5811_EQUIPMENT	0.0101	0.0007
⋮	⋮	⋮	⋮

Table 4.2 K-W test results by yield and bin 4 item

		yield		
		significance	Non-significance	
bin 4 item	significance	99	144	243
	Non-significance	59	489	548
		158	633	791

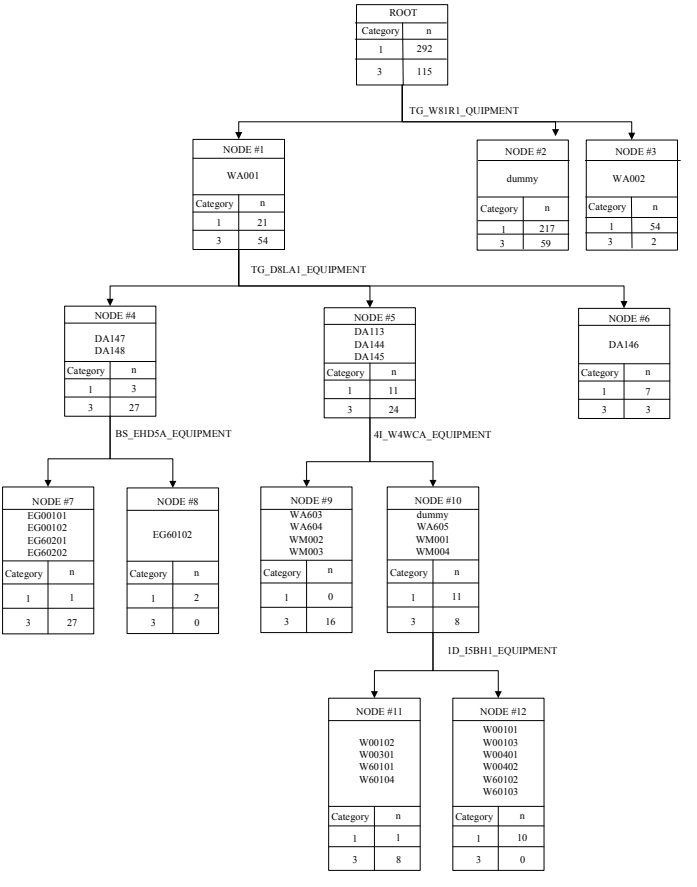


Fig. 4.3 The result of MBDT

4.2 Results and Discussion

The RP threshold and the FCSIG threshold are set as 1. Cluster#3 is the FC. The tree conducted by MBDT algorithm is shown as Fig. 4.3. The derived rules can provide reference for engineers to quickly identify abnormal machine and remove

the assignable causes. To consider the rules for Cluster#3, the equipment WA001 in the process TG_W81R1 is the first level of MDBT could be in serious abnormal condition. In addition, the equipments DA147 and DA148 in the process TG_D8LA1_EQ, and equipments WA604, WA604, WM002, and WM003 in the process 4I_W4WCA_EQ are also the potential root causes from the MDBT rules.

Although, the equipment have been maintained, it is still necessary to inspect the current operation behavior and continuously monitor for ensuring operation well. According to the discussion with engineers, we compare the daily inspection data and daily work report in production to trace the low yield pattern. Operators and mechanical engineers prove that this abnormal condition of equipment caused the bin 4 item becoming high was occurred after the prevent maintenance (PM).

5 Conclusion

The proposed MBDT algorithm can effectively diagnose the root causes and provide quick response to abnormal equipment in the complex semiconductor manufacturing processes with imbalanced classes. Indeed, the results of empirical study have validated the practical viability of the proposed approach. The proposed MI approach can help engineers identify the hidden information from huge database and relationship between manufacturing data and off-line test data. The proposed approach can construct smaller but more interpretable trees structure.

From the data mining methods to the results presentation, including the numerical data, visualization, or rules description, analyzer should interact with domain experts for their experience and advice. Through the repeatedly circle of analysis and discussion, mining results will be more helpful and effective to engineers. Furthermore, the semiconductor mining framework for yield enhancement can help building the decision support mechanism of fault diagnosis. Further study can be done to construct a system based on the proposed framework, in which the embedded MI methods can be applied to advanced equipment control and advanced process control.

Acknowledgments. This research is supported by National Science Council, Taiwan (NSC99-2221-E-007-047-MY3; NSC 100-2410-H-155-048).

References

[1] Berry, M., Linoff, G.: Data Mining Techniques for Marketing, Sales and Customer Support, 2nd edn. John Wiley and Sons, New York (2004)

[2] Braha, D., Shmilovici, A.: Data mining for improving a cleaning process in the semiconductor industry. IEEE Trans. Semicond. Manuf. 15(1), 91–101 (2002)

[3] Chiang, C., Kawa, J.: Design for Manufacturability and Yield for Nano-scale CMOS. Springer, The Netherlands (2007)

[4] Chien, C.-F., Hsu, C.: A novel method for determining machine subgroups and backups with an empirical study for semiconductor manufacturing. J. Intell. Manuf. 17(4), 429–440 (2006)

[5] Chien, C.-F., Wang, W., Cheng, J.: Data mining for yield enhancement in semiconductor manufacturing and an empirical study. Expert Syst. Appl. 33(1), 192–198 (2007)
[6] Chien, C., Chen, Y., Peng, J.: Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and production life cycle. Int. J. Prod. Econ. 128, 496–509 (2010)
[7] Chien, C.-F., Hsu, C., Hsiao, C.: Manufacturing intelligence to forecast and reduce semiconductor cycle time. J. Intell. Manuf. (2011), doi:10.1007/s10845-011-0572-y
[8] Hsu, C., Chien, C.-F., Lin, K., Chien, C.: Data mining for yield enhancement in TFT-LCD manufacturing and an empirical study. Journal of the Chinese Institute of Industrial Engineers 27(2), 140–156 (2010)
[9] Hsu, S., Chien, C.-F.: Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. Int. J. Prod. Econ. 107, 88–103 (2007)
[10] Japkowicz, N., Shaju, S.: The class imbalance problem: a systematic study. Intell. Data. Anal. 6(5), 429–450 (2002)
[11] Kotsiantis, S.B., Pintelas, P.E.: Mixture of expert agents for handling imbalanced data sets. Ann. Math Comput. Teleinformatics 1(1), 46–55 (2003)
[12] Kuo, C., Chien, C.-F., Chen, C.: Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time. IEEE Trans. Autom. Sci. Eng. 8(1), 103–111 (2010)
[13] Weiss, G.: Mining with rarity: A unifying framework. SIGKDD Explor. 6(1), 7–19 (2004)
[14] Zhou, C., Nelson, P.C., Xiao, W., Tripak, T.M., Lane, S.A.: An intelligent data mining system for drop test analysis of electronic products. IEEE Trans. Electron. Packag. Manuf. 24(3), 222–231 (2001)

Medical Discriminant Analysis by Modular Fuzzy Model

Hirosato Seki and Toshihiko Watanabe

Abstract. Since the number of rules of the single input rule modules connected fuzzy inference method (SIRMs method) is limited as compared to the conventional fuzzy inference method, inference results gained by the SIRMs method are simple in general. From this reason, Watanabe et al. have proposed a *modular fuzzy model* which extends the SIRMs method, and shown that this model can obtain good results for reinforcement learning. In this paper, this model is compared with the conventional fuzzy inference method by applying to discriminant analysis of a medical data.

1 Introduction

As for the "IF-THEN" rules in the conventional fuzzy inference methods [5], all the input items of the system are set to the antecedent part, and all output items are set to the consequent part. Therefore, the problem is apparent that the number of fuzzy rules becomes increasingly huge; hence, the setup and adjustment of fuzzy rules become difficult. On the other hand, the *single input rule modules connected type fuzzy inference method* (SIRMs method) by Yubazaki et al. [17, 18, 19, 20, 21, 23] which unifies the inference output from fuzzy rule modules of one input type "IF-THEN" form can reduce the number of fuzzy rules drastically. The method has been applied to nonlinear function identification, control of a first order lag system with dead time, orbital pursuit control of a non-restrained object, and stabilization control of a handstand system etc., and good results are obtained. However, since

Hirosato Seki
Kwansei Gakuin University, 2-1, Gakuen, Sanda-shi, Hyogo, 669-1337, Japan
e-mail: h_seki@ieee.org

Toshihiko Watanabe
Osaka Electro-Communication University, 18-8, Hatsu-cho, Neyagawa-shi, Osaka, 572-8530, Japan
e-mail: t-wata@isc.osakac.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 245–251.
springerlink.com

the number of rules of the SIRMs method is limited compared to the conventional inference methods, inference results gained by the SIRMs method are simple in general. Especially, Seki et al have shown that the SIRMs method can not be applied to the realization of exclusive OR (XOR) [6, 7].

From these reasons, Watanabe et al. have proposed the *modular fuzzy model* extending SIRMs method [13, 14, 15, 16]. However, the applicability of the modular fuzzy model is not clarified in real systems.

Therefore, in this paper, we apply the modular fuzzy model to a medical diagnosis, and it is compared with the simplified fuzzy inference method as the conventional method.

2 Modular Fuzzy Model

2.1 Structure of Modular Fuzzy Model

As a fuzzy model having high applicability, Single Input Rule Modules(SIRMs) [17, 18, 19, 20, 21, 23] have been proposed. The idea is to unify reasoning outputs from fuzzy rule modules comprised with single input and single output formed fuzzy if-then rules. The number of rules can be drastically reduced as well as bringing us high maintainability in actual applications. However, its disadvantage of low precision is inevitable in order to apply the method to high dimensional or complicated problems. The modular fuzzy model [13, 14, 15, 16] is extension of the SIRMs method by relaxing the restriction of the input space, i.e. single, to arbitrary subspace of the rule for constructing the model of huge multi-dimensional space. In many application cases of large multi-dimensional space, human expertise or know-how can almost be expressed approximately using some inputs among whole input variables, e.g. fuzzy control and expert system. The concept of the modular fuzzy model is intuitively representing such observations. Description of the model is as follows:

$$\begin{aligned}
&\text{Rules-1} : \{P_1(x) = A_j^1 \longrightarrow y_1 = f_j^1(P_1(x))\}_{j=1}^{m_1} \\
&\qquad\qquad \vdots \\
&\text{Rules-}i : \{P_i(x) = A_j^i \longrightarrow y_i = f_j^i(P_i(x))\}_{j=1}^{m_i} \\
&\qquad\qquad \vdots \\
&\text{Rules-}n : \{P_n(x) = A_j^n \longrightarrow y_n = f_j^n(P_n(x))\}_{j=1}^{m_n}
\end{aligned} \tag{1}$$

where "Rules-i" stands for the ith fuzzy rule module, $P_i(x)$ denotes predetermined projection of the input vector x in the ith module, y_i is the output variable, and n is the number of rule modules. The number of constituent rules in the ith fuzzy rule module is m_i. f is the function of consequent part of the rule like T–S inference model [11]. A_j^i denotes the fuzzy sets defined in the projected subspace. The membership degree of the antecedent part of the jth rule in the "Rules-i" module is calculated as:

$$h_j^i = A_j^i(P_i(x^0)) \tag{2}$$

where h_j^i denotes the membership degree, and x^0 is an input vector. The output of fuzzy reasoning of each module is decided as the following equation.

$$y_i^0 = \frac{\sum_{j=1}^{m_i} h_j^i f_j^i(P_i(x^0))}{\sum_{j=1}^{m_i} h_j^i} \tag{3}$$

The final output of the Modular Fuzzy Model is formulated as:

$$y^0 = \sum_{i=1}^{n} w_i y_i^0 \tag{4}$$

2.2 Definition of Projection

In order to construct the modular fuzzy model, the necessary projection "P( )" of each "Rule Module" should be defined. We explain the basic idea of the definition using examples. Assume that x is 4 dimensional input vector as:

$$x = [x_1, x_2, x_3, x_4]^T \tag{5}$$

Then, the conventional SIRMs architecture adopts the projection as follows:

$$\{P_1(x), P_2(x), P_3(x), P_4(x)\} = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}\} \tag{6}$$

Though the projection in the modular fuzzy model can be arbitrarily defined, the primary idea of the modular fuzzy model is to define the projection from combination of each variable as:

$$\begin{aligned} &\{P_1(x), P_2(x), P_3(x), P_4(x), P_5(x), P_6(x)\} = \\ &\{(x_1, x_2), (x_1, x_3), (x_1, x_4), (x_2, x_3), (x_2, x_4), (x_3, x_4)\} \end{aligned} \tag{7}$$

where pair inputs are used for rule modules. The modular structure can be formulated from the definition of the number of projection dimension. The number of projection dimension is 2 in this case. In the same way, triplet inputs, i.e. 3-dimensional projection, can also be formulated as:

$$\begin{aligned} &\{P_1(x), P_2(x), P_3(x), P_4(x)\} = \\ &\{(x_1, x_2, x_3), (x_1, x_2, x_4), (x_1, x_3, x_4), (x_2, x_3, x_4)\} \end{aligned} \tag{8}$$

In this manner, the construction of projection leads to reduction of the number of total rules. The number of total rules in modular fuzzy model can be less than the

conventional Cartesian product typed fuzzy model when the number of modules is limited. Furthermore, projections for the rule modules can be selected from the whole combination, i.e. power set of attribute items in the input vector. We can also utilize the same projection for plural modules in different definition of fuzzy partition. Though the problem of deciding the projection dimension and the appropriate number of rule modules still remains, it is decided by increasing step by step through modeling from one dimension, i.e. equivalent to SIRMs, to the arbitrary number of projection dimension in this study.

3 Learning Algorithms for Modular Fuzzy Model

The setup of membership functions and fuzzy rules is difficult when systems involute. Hence, we expect to automatically optimize membership functions and fuzzy rules based on input-output data in systems. From the reason, the learning algorithms for membership functions and fuzzy rules were proposed [2, 3, 4, 9, 12, 22]. In this section, we review a learning algorithm of the modular fuzzy model. As for the learning method, the parameters are learned for the membership functions of the antecedent parts, consequent values, and importance degrees of input items.

When the training input–output data $(x_1, x_2, \ldots, x_n; y^T)$ are given for a fuzzy system model, it is usual to use the following objective function E for evaluating an error between y^T and y^0, which can be regarded as an optimum problem:

$$E = \frac{1}{2}(y^T - y^0)^2 \tag{9}$$

where y^T is the desired output value, and y^0 the corresponding fuzzy inference result.

The Gaussian-type fuzzy sets are used as antecedent parts and cnsequent parts. The parameters of center, width, consequent value f_j^i and importance degree w_i are obtained by the steepest descent method. The consequent value f_j^i's are used as constant in this case.

4 Formulation of a Medical Diagnosis System by Modular Fuzzy Model

In this section, the modular fuzzy model is applied to a medical diagnosis, and compared with the neuro-fuzzy method based on the simplified fuzzy inference method. 145 real diabetes data are used [1]. The data have 5 input items and 1 output, and the 5 input items are constituted as follows.

- Relative weight
- Fasting levels of plasma glucose concentrations
- Glucose value
- Insulin value
- SSPG (Steady State Plasma Glucose)

The inputs $x_1, x_2, \ldots, x_5$ are normalized as in [0, 1]. The data is classified Group 1 (Clinical Diabetes), Group 2 (Chemical Diabetes) or Group 3 (Soundness). The Group 1, Group 2 and Group 3 in a desired output which uses to inference are 1, 0.5 and 0, respectively. Moreover, inference result y^0 in fuzzy infernce is classified as follows.

$$\begin{cases} 1\ (\text{ClinicalDiabetes}) & : y^0 \geq 0.75 \\ 0.5\ (\text{ChemicalDiabetes}) & : 0.25 \leq y^0 < 0.75 \\ 0\ (\text{Soundness}) & : 0.25 < y^0 \end{cases}$$

In medical diagnosis, we use four membership functions for five inputs $x_1, x_2, \ldots, x_5$, where the centers of the membership functions $A_1^i, A_2^i, A_3^i, A_4^i$ for $i = 1, 2, \ldots, 5$ are -0.5, 0, 0.5, 1, and each width of the membership functions is 0.5. The number of combination for input is 2, i.e., the number of input is extended to 10 dimensions in modular fuzzy model. Moreover, all of the consequent parts and importance degree for each input item are set to be 0 and 0.5, respectively.

In our case, 72 training data are employed from 145 diabetes data in random order, and 73 checking data are used from the remaining diabetes data.

In the following, we apply the modular fuzzy model to medical data in the case of using the Gaussian-type membership functions. Moreover, this method is also compared with the neuro-fuzzy based on the simplifid fuzzy inference method as the conventional method.

For the above, learninig iterations are executed 1000 times, and 10 simulations are run. Table 1 shows the answer ratio using the checking data for medical diagnosis, where NF stands for the neuro-fuzzy based on the simplifidfuzzy inference method in the table. Table 2 shows the error of evaluation using the checking data for medical diagnosis, where the error of evaluation denotes a mean square error for the checking data.

Although the total number of rules of the neuro-fuzzy method uses 1024 ($= 4^5$), that of modular fuzzy model are 40 ($=_5 C_2 \times 4$) rules. Even as the modular fuzzy

Table 1 Answer ratio(%) for medical data

Case	Modular	NF [9]
1	72.6	69.9
2	72.6	68.5
3	76.7	72.6
4	65.8	72.6
5	71.2	71.2
6	67.1	72.6
7	65.8	61.6
8	75.3	82.2
9	69.9	72.6
10	76.7	80.8
Average	71.4	72.5

Table 2 Error evaluation for medical data

Case	Modular	NF [9]
1	0.05303	0.06138
2	0.05183	0.06357
3	0.07191	0.05496
4	0.07501	0.06360
5	0.06266	0.06238
6	0.05567	0.05674
7	0.07435	0.07089
8	0.04848	0.03745
9	0.05516	0.05938
10	0.07874	0.04297
Average	0.06268	0.05733

model uses few rules, the inference results obtained by the modular fuzzy model are coequal to the neuro-fuzzy method.

From the reason, the modular fuzzy model will be useful compared with neuro-fuzzy method in the medical data.

5 Conclusions

In this paper, we have shown the applicability of the modular fuzzy model by using a medical diagnosis. Even as the modular fuzzy model uses few rules, the modular fuzzy model can be obtained good results. Therefore, the modular fuzzy model will be used as simple model in real systems.

Further studies are required to optimize input combination, and select from all input mapped to high dimension.

References

1. Andrews, D.F., Herzberg, A.M.: Data:A collection of problems from many fields for the students and research worker. Springer (1985)
2. Cui, X., Shin, K.G.: Direct control and coordination using neural networks. IEEE Trans. Syst., Man, Cybern. 23, 686–697 (1993)
3. Ichihashi, H.: Iterative fuzzy modeling and a hierarchical network. In: Proc. the 4th IFSA Congress of Engineering, pp. 49–52 (1991)
4. Juang, C.F.: A TSK type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms. IEEE Trans. Fuzzy Systems 10, 155–170 (2002)
5. Mamdani, E.H.: Application of fuzzy algorithms for control of simple dynamic plant. Proc. IEEE 121, 1585–1588 (1974)
6. Seki, H., Ishii, H., Mizumoto, M.: On the property of single input rule modules connected type fuzy reasoning method. In: 2007 IEEE International Conference on Fuzzy Systems, pp. 1180–1185 (2007)

7. Seki, H., Ishii, H., Mizumoto, M.: On the generalization of single input rule modules connected type fuzy reasoning method. IEEE Trans. Fuzzy Systems 16, 1180–1187 (2008)
8. Seki, H., Ishii, H., Mizumoto, M.: On the monotonicity of fuzzy-inference methods related to T–S inference method. IEEE Trans. Fuzzy Systems 18, 629–634 (2010)
9. Seki, H., Ishii, H., Mizumoto, M.: Nonlinear identification and medical diagnosis system using functional-type SIRMs connected fuzzy inference method. International Journal of Innovative Computing, Information and Control 6, 5275–5286 (2010)
10. Seki, H., Mizumoto, M.: On the equivalence conditions of fuzzy inference methods—part 1: basic concept and definition. IEEE Trans. Fuzzy Systems 19, 1097–1106 (2011)
11. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its Applications to modeling and control. IEEE Trans. Syst., Man, Cybern. 15, 116–132 (1985)
12. Wang, L.X., Mendel, J.M.: Back-Propagation fuzzy system as nonlinear dynamic system identifiers. In: Proc. 1992 IEEE Int. Conf. Fuzzy Systems, pp. 1409–1416 (1992)
13. Watanabe, T., Takahashi, Y.: Hierarchical reinforcement learning using a modular fuzzy model for multi-agent problem. In: Proc. 2007 IEEE Int. Conf. Syst., Man, Cybern., pp. 1681–1686 (2007)
14. Watanabe, T., Fujioka, R.: Collaborative filtering based on modular fuzzy model concept. In: Proc. Joint 4th Int. Conf. on Soft Computing and Intelligent Systems and 9th Int. Symp. on Advanced Intelligent Systems (SCIS & ISIS 2008), pp. 1350–1354 (2008)
15. Watanabe, T., Wada, T.: Reinforcement learning based on modular fuzzy model with gating unit. In: Proc. 2008 IEEE Int. Conf. Syst., Man, Cybern., pp. 1806–1811 (2008)
16. Watanabe, T., Seki, H.: A study on modeling approach based on modular fuzzy model. In: Proc. The 1st International Workshop on Aware Computing (IWAC 2009), pp. 743–748 (2009)
17. Yi, J., Yubazaki, N., Hirota, K.: Stabilization control of seriestype double inverted pendulum systems using the SIRMs dynamically connected fuzzy inference model. Artificial Intelligence in Engineering 15, 297–308 (2001)
18. Yi, J., Yubazaki, N., Hirota, K.: Upswing and stabilization control of inverted pendulum system based on the SIRMs dynamically connected fuzzy inference model. Fuzzy Sets and Systems 122, 139–152 (2001)
19. Yi, J., Yubazaki, N., Hirota, K.: A new fuzzy controller for stabilization of parallel-type double inverted pendulum system. Fuzzy Sets and Systems 126, 105–119 (2002)
20. Yi, J., Yubazaki, N., Hirota, K.: A proposal of SIRMs dynamically connected fuzzy inference model for plural input fuzzy control. Fuzzy Sets and Systems 125, 79–92 (2002)
21. Yi, J., Yubazaki, N., Hirota, K.: Anti-swing and positioning control of overhead traveling crane. Inf. Sci. 155, 19–42 (2002)
22. Yu, W., Li, X.: Fuzzy identification using fuzzy neural networks with stable learning algorithms. IEEE Trans. Fuzzy Systems 12, 411–420 (2004)
23. Yubazaki, N., Yi, J., Hirota, K.: SIRMs (Single Input Rule Modules) connected fuzzy inference model. Journal of Advanced Computational Intelligence and Intelligent Informatics 1, 23–30 (1997)

Normalization Method Based on Dummy Alternative with Perfect Evaluation Score in AHP and ANP

Kazutomo Nishizawa

Abstract. In this paper an appropriate normalization method for alternatives in the Analytic Hierarchy Process (AHP) and the Analytic Network Process (ANP) is proposed. In the traditional AHP, in general, we calculate the principal eigenvalue and the corresponding eigenvector from a pairwise comparison matrix. By normalizing the sum of eigenvector to 1, we obtain the weights of alternatives. However, this normalization method is considered as one of the causes of some problems, for example rank reversal etc.. Therefore, at first in this paper, the verification example for verify the normalization methods is shown and also the correct value of this example is shown. Using this example, at secondly, some normalization problems are pointed out. Based on the incorrect results, at thirdly, an appropriate normalization method based on artificially added dummy alternative with the perfect evaluation score is proposed. Applying proposed method to the verification example for the AHP and the two-cluster ANP, the accurate result is obtained. Finally, proposed normalization method is concluded by considering the result from the verification example.

1 Introduction

In this paper an appropriate normalization method for alternatives in the Analytic Hierarchy Process (AHP) [7] and the Analytic Network Process (ANP) [8] is proposed.

In the traditional AHP, in general, we calculate the principal eigenvalue and the corresponding eigenvector from a pairwise comparison matrix. By normalizing the sum of eigenvector to 1, we obtain the weights of alternatives. However, this normalization method is considered as one of the cause of some

Kazutomo Nishizawa
Nihon University, 1-2-1 Izumicho, Narashino, Chiba 275-8575, Japan
e-mail: nishizawa.kazutomo@nihon-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 253–262.
springerlink.com

problems, for example rank reversal etc.. This problem was pointed out by Belton and Gear and was discussed [1, 2]. Then, some improved methods were proposed. For example, the Linking Pin AHP was proposed by Schoner and Wedley [9, 10] and the Dominant AHP was proposed by Kinoshita and Nakanishi [3]. These improved AHP models were compared and discussed by Kinoshita and Sugiura [4].

On the other hand, the two-cluster ANP, equivalent to the AHP, was proposed on our previous study [5, 6]. In this method, the improved supermatrix which consists of the sub-matrix, based on the weights of alternatives, and its transposed matrix was proposed.

However these methods, mentioned above, were not improved normalization of the weights of alternatives. Therefore, at first in this paper, the verification example for verify the normalization methods is shown and also the correct value of this example is shown. Using this example, some normalization problems are pointed out. Based on the incorrect results of calculation, an appropriate normalization method based on artificially added dummy alternative with the perfect evaluation score is proposed. Applying proposed method to the verification example for the AHP and the two-cluster ANP, the accurate result is obtained.

This paper consists of following sections. At first, the verification example for verify the normalization methods is shown in section 2. In section 3, the normalization method in the AHP is proposed, and apply to the verification example, shown in section 2. Next, proposed method is applied to the two-cluster ANP in section 4. Finally, In section 5, proposed normalization method is concluded by considering the result from the verification example.

2 Verification Example for Verify the Normalization Methods

In this section, the verification example for verify the normalization methods is shown and also the correct value of this example is shown. Table 1 is the verification example. This is the example of students assessment.

Table 1 Verification example (students assessment)

	c_1	c_2	c_3	c_4	Total score
	20	20	10	50	
a_1	20	12	10	35	77
a_2	20	18	10	44	92
a_3	19	12	9	4	44
a_4	16	13	7	26	62
a_5	20	17	10	34	81

In this example, five students ($a_1 \sim a_5$) were evaluated by the four criteria ($c_1 \sim c_4$). In Table 1, a symbol of the criterion c_1 is an exercise, c_2 is a homework, c_3 is a report, and c_4 is an examination. The allotment score of criteria ($c_1 \sim c_4$), in this example, is 20, 20, 10 and 50, respectively. The final evaluation of the students are performed by the total score.

The matrix form of the total score calculation is given by Eq. (1).

$$\begin{bmatrix} 20 & 12 & 10 & 35 \\ 20 & 18 & 10 & 44 \\ 19 & 12 & 9 & 4 \\ 16 & 13 & 7 & 26 \\ 20 & 17 & 10 & 34 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 77 \\ 92 \\ 44 \\ 62 \\ 81 \end{bmatrix} \tag{1}$$

In the AHP, in actual cases, we need appropriate normalization for the weights of alternatives. The form of the calculation is shown in Eq. (2).

$$\mathbf{W} \times \mathbf{v} = \mathbf{w}, \tag{2}$$

where a matrix $\mathbf{W}$ is the weights of alternatives which were evaluated by each criteria, and a vector $\mathbf{v}$ is the weight of criteria. A vector $\mathbf{w}$ is the final weight of alternatives.

Therefore, normalizing verification example by the allotment score, Table 2 is obtained.

Table 2 Normalized verification example by the allotment score

	c_1	c_2	c_3	c_4	Total score
	20	20	10	50	
a_1	1.000	0.600	1.000	0.700	77
a_2	1.000	0.900	1.000	0.880	92
a_3	0.950	0.600	0.900	0.080	44
a_4	0.800	0.650	0.700	0.520	62
a_5	1.000	0.850	1.000	0.680	81

Based on Eq. (2), the AHP form of Table 2 is given by Eq. (3).

$$\begin{bmatrix} 1.000 & 0.600 & 1.000 & 0.700 \\ 1.000 & 0.900 & 1.000 & 0.880 \\ 0.950 & 0.600 & 0.900 & 0.080 \\ 0.800 & 0.650 & 0.700 & 0.520 \\ 1.000 & 0.850 & 1.000 & 0.680 \end{bmatrix} \begin{bmatrix} 20 \\ 20 \\ 10 \\ 50 \end{bmatrix} = \begin{bmatrix} 77 \\ 92 \\ 44 \\ 62 \\ 81 \end{bmatrix} \tag{3}$$

Furthermore, normalizing the sum of the allotment scores to 1, Table 3 is obtained.

Table 3 Normalized verification example by the sum of the allotment scores to 1

	c_1	c_2	c_3	c_4	
	0.20	0.20	0.10	0.50	
a_1	1.000	0.600	1.000	0.700	0.770
a_2	1.000	0.900	1.000	0.880	0.920
a_3	0.950	0.600	0.900	0.080	0.440
a_4	0.800	0.650	0.700	0.520	0.620
a_5	1.000	0.850	1.000	0.680	0.810

The AHP form of Table 3 is given by Eq. (4).

$$\begin{bmatrix} 1.000 & 0.600 & 1.000 & 0.700 \\ 1.000 & 0.900 & 1.000 & 0.880 \\ 0.950 & 0.600 & 0.900 & 0.080 \\ 0.800 & 0.650 & 0.700 & 0.520 \\ 1.000 & 0.850 & 1.000 & 0.680 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.20 \\ 0.10 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.770 \\ 0.920 \\ 0.440 \\ 0.620 \\ 0.810 \end{bmatrix} \tag{4}$$

Normalizing the sum of the result (4) to 1, the weight of alternatives $\mathbf{w}$ is obtained.

$$\mathbf{w} = \begin{bmatrix} 0.216292 \\ 0.258427 \\ 0.123596 \\ 0.174157 \\ 0.227528 \end{bmatrix} \tag{5}$$

Obtained (5) is the correct values of the verification example of Table 1. In this paper, normalization methods are verified by these values.

3 Proposed Normalization Method in the AHP

Based on the incorrect results of normalization, in this section, an appropriate normalization method for the alternatives in the AHP is proposed.

In the traditional AHP [7], we normalized the sum of the weights of alternatives for each criterion to 1. Applying such method to example in Table 1, Table 4 is obtained.

The AHP form of Table 4 is given by Eq. (6).

$$\begin{bmatrix} 0.211 & 0.167 & 0.217 & 0.245 \\ 0.211 & 0.250 & 0.217 & 0.308 \\ 0.200 & 0.167 & 0.196 & 0.028 \\ 0.168 & 0.181 & 0.152 & 0.182 \\ 0.211 & 0.236 & 0.217 & 0.238 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.20 \\ 0.10 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.219555 \\ 0.267691 \\ 0.106885 \\ 0.175922 \\ 0.229948 \end{bmatrix} \tag{6}$$

Table 4 Normalized verification example by the traditional AHP procedure

	c_1	c_2	c_3	c_4
	0.20	0.20	0.10	0.50
a_1	0.211	0.167	0.217	0.245
a_2	0.211	0.250	0.217	0.308
a_3	0.200	0.167	0.196	0.028
a_4	0.168	0.181	0.152	0.182
a_5	0.211	0.236	0.217	0.238

Unfortunately, the result (6) is not coincides with the correct values (5). Therefore, this result reveals that the traditional AHP procedure, normalizing the sum of the weights of alternatives for each criterion to 1, is not appropriate.

On the other hand, in previous study [6], some better results were obtained by normalizing the maximum weight of alternative to 1. The normalized results of alternatives are shown in Table 5.

Table 5 Normalized verification example by maximum alternative value to 1

	c_1	c_2	c_3	c_4
	20	20	10	50
a_1	1.000	0.667	1.000	0.795
a_2	1.000	1.000	1.000	1.000
a_3	0.950	0.667	0.900	0.091
a_4	0.800	0.722	0.700	0.591
a_5	1.000	0.944	1.000	0.773

The AHP form of Table 5 is given by Eq. (7).

$$\begin{bmatrix} 1.000 & 0.667 & 1.000 & 0.795 \\ 1.000 & 1.000 & 1.000 & 1.000 \\ 0.950 & 0.667 & 0.900 & 0.091 \\ 0.800 & 0.722 & 0.700 & 0.591 \\ 1.000 & 0.944 & 1.000 & 0.773 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.20 \\ 0.10 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.831 \\ 1.000 \\ 0.459 \\ 0.670 \\ 0.875 \end{bmatrix} \tag{7}$$

Normalizing the sum of the result (7) to 1, the weight of alternatives $\mathbf{w}$ is obtained.

$$\mathbf{w} = \begin{bmatrix} 0.216704 \\ 0.260756 \\ 0.119632 \\ 0.174680 \\ 0.228228 \end{bmatrix} \tag{8}$$

The result (8) is also not coincides with the correct values (5). However, Eq.(6) and Eq.(8) are suggest us some hints of appropriate normalization method.

With respect to criterion c_2 in Table 5, despite there are no perfect evaluation score in Table 1, the value of a_2 is 1. With respect to c_4 in Table 5, the value of a_2 is also similar. It is clear that we need to divide each alternative value into allotment score, or perfect evaluation score.

The procedure of proposed normalization method as follows.

1. Add dummy alternative with perfect evaluation score to the comparison matrix.
2. Calculate the principal eigenvalue and the corresponding eigenvector of this comparison matrix.
3. Divide each eigenvector into eigenvector of dummy alternative.

According to the above procedure, add dummy alternative d with perfect evaluation score to Table 1, Table 6 is obtained. However, in this paper, calculation process from the comparison matrix is omitted.

Table 6 Add dummy alternative with perfect evaluation score

	c_1	c_2	c_3	c_4	Total score
	20	20	10	50	
a_1	20	12	10	35	77
a_2	20	18	10	44	92
a_3	19	12	9	4	44
a_4	16	13	7	26	62
a_5	20	17	10	34	81
d	20	20	10	50	100

Normalizing each alternatives score by the perfect evaluation score of d, Table 7 is obtained.

By normalized the sum of criteria weights to 1, the AHP form of Table 7 is given by Eq. (9).

$$\begin{bmatrix} 1.000 & 0.600 & 1.000 & 0.700 \\ 1.000 & 0.900 & 1.000 & 0.880 \\ 0.950 & 0.600 & 0.900 & 0.080 \\ 0.800 & 0.650 & 0.700 & 0.520 \\ 1.000 & 0.850 & 1.000 & 0.680 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.20 \\ 0.10 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.770 \\ 0.920 \\ 0.440 \\ 0.620 \\ 0.810 \end{bmatrix} \tag{9}$$

Normalizing the sum of (9) equal to 1, the result is coincides with the correct values (5).

Table 7 Normalized by the dummy alternative

	c_1	c_2	c_3	c_4	Total score
	20	20	10	50	
a_1	1.000	0.600	1.000	0.700	77
a_2	1.000	0.900	1.000	0.880	92
a_3	0.950	0.600	0.900	0.080	44
a_4	0.800	0.650	0.700	0.520	62
a_5	1.000	0.850	1.000	0.680	81
d	1.000	1.000	1.000	1.000	100

Based on the calculation results of the AHP, an appropriate normalization method for the ANP is considered in next section.

4 Apply Proposed Method to the Two-Cluster ANP

In this section, proposed normalization method, described in section 3, is applied to the ANP.

In the traditional ANP [8], we construct the supermatrix $\mathbf{S}$ and obtain the converged values of $\mathbf{S}^{\infty}$, where $\mathbf{S}$ is a stochastic matrix. From the result of converged matrix, we can obtain the weights of criteria and the weights of alternatives. Therefore the normalization of the two-cluster ANP which equivalent to the AHP is considered.

The form of the supermatrix of this case is shown in Eq. (10), where a sub-matrix $\mathbf{W}$ is an evaluation result of the alternatives, a sub-matrix $\mathbf{V}$ is an evaluation result of the criteria and $\mathbf{0}$ is a zero matrix. In the ANP, a sub-matrix $\mathbf{W}$ is constructed in the similar procedure to the AHP. On the other hand, in the AHP, we do not construct a sub-matrix $\mathbf{V}$.

$$\mathbf{S} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{V} \\ \hline \mathbf{W} & \mathbf{0} \end{array}\right] \tag{10}$$

In this case we can represent the weights of alternatives $\mathbf{w} = \mathbf{W} \times \mathbf{v}$ and the weights of criteria $\mathbf{v} = \mathbf{V} \times \mathbf{w}$, therefor Eq. (11) is obtained.

$$\begin{bmatrix} \mathbf{0} & \mathbf{V} \\ \mathbf{W} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \tag{11}$$

Note that Eq. (11) is an eigenvalue problem of $\mathbf{S}$ with eigenvalue 1. Therefore we can obtain the weights instead of to calculate $\mathbf{S}^{\infty}$.

On the other hand, in previous study [5, 6], the improved supermatrix which consists of the sub-matrix $\mathbf{W}$ and the transposed sub-matrix $\mathbf{W}^T$ was

proposed. In this method, instead of $\mathbf{V}$, $\mathbf{W}^T$ was introduced. The form of the improved supermatrix is shown in Eq. (12).

$$\mathbf{S} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{W}^T \\ \hline \mathbf{W} & \mathbf{0} \end{array}\right] \tag{12}$$

Now based on Table 1, the improved supermatrix $\mathbf{S_0}$ is constructed in Eq. (13).

$$\mathbf{S_0} = \left[\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 20 & 20 & 19 & 16 & 20 \\ 0 & 0 & 0 & 0 & 12 & 18 & 12 & 13 & 17 \\ 0 & 0 & 0 & 0 & 10 & 10 & 9 & 7 & 10 \\ 0 & 0 & 0 & 0 & 35 & 44 & 4 & 26 & 34 \\ \hline 20 & 12 & 10 & 35 & 0 & 0 & 0 & 0 & 0 \\ 20 & 18 & 10 & 44 & 0 & 0 & 0 & 0 & 0 \\ 19 & 12 & 9 & 4 & 0 & 0 & 0 & 0 & 0 \\ 16 & 13 & 7 & 26 & 0 & 0 & 0 & 0 & 0 \\ 20 & 17 & 10 & 34 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{13}$$

In this way, the improved supermatrix (13) is obtained, however, $\mathbf{S_0}$ is not stochastic matrix. Transform $\mathbf{S_0}$ to stochastic matrix $\mathbf{S}$ and the eigenvector of $\mathbf{S}$ is calculated. Further normalizing the sum of the eigenvector equal for each criterion to 1, the weights of alternatives (14) is obtained.

$$\mathbf{w} = \begin{bmatrix} 0.216292 \\ 0.258427 \\ 0.123596 \\ 0.174157 \\ 0.227528 \end{bmatrix} \tag{14}$$

The result (14) is coincide with the correct values (5).

Next, the supermatrix is constructed by the proposed normalization method described in Section 3. Based on Table 7, the supermatrix is given by Eq. (15).

$$\mathbf{S_0} = \left[\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 1.000 & 1.000 & 0.950 & 0.800 & 1.000 \\ 0 & 0 & 0 & 0 & 0.600 & 0.900 & 0.600 & 0.650 & 0.850 \\ 0 & 0 & 0 & 0 & 1.000 & 1.000 & 0.900 & 0.700 & 1.000 \\ 0 & 0 & 0 & 0 & 0.700 & 0.880 & 0.080 & 0.520 & 0.680 \\ \hline 1.000 & 0.600 & 1.000 & 0.700 & 0 & 0 & 0 & 0 & 0 \\ 1.000 & 0.900 & 1.000 & 0.880 & 0 & 0 & 0 & 0 & 0 \\ 0.950 & 0.600 & 0.900 & 0.080 & 0 & 0 & 0 & 0 & 0 \\ 0.800 & 0.650 & 0.700 & 0.520 & 0 & 0 & 0 & 0 & 0 \\ 1.000 & 0.850 & 1.000 & 0.680 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{15}$$

Transform (15) to stochastic matrix, the weights of alternatives (16) is obtained.

$$\mathbf{w} = \begin{bmatrix} 0.208729 \\ 0.239089 \\ 0.160025 \\ 0.168880 \\ 0.223276 \end{bmatrix} \tag{16}$$

Unfortunately, the result (16) is not coincide with the correct values (5). It is clear that normalized sub-matrix $\mathbf{W}$ is not including normalization information. Therefore, multiply the normalized allotment scores in Table 3 to $\mathbf{W}$, as a result, improved supermatrix (17) is constructed and the weight of alternatives (18) is obtained.

$$\mathbf{S_0} = \left[\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 0.200 & 0.200 & 0.190 & 0.160 & 0.200 \\ 0 & 0 & 0 & 0 & 0.120 & 0.180 & 0.120 & 0.130 & 0.170 \\ 0 & 0 & 0 & 0 & 0.100 & 0.100 & 0.090 & 0.070 & 0.100 \\ 0 & 0 & 0 & 0 & 0.350 & 0.440 & 0.040 & 0.260 & 0.340 \\ \hline 0.200 & 0.120 & 0.100 & 0.350 & 0 & 0 & 0 & 0 & 0 \\ 0.200 & 0.180 & 0.100 & 0.440 & 0 & 0 & 0 & 0 & 0 \\ 0.190 & 0.120 & 0.090 & 0.040 & 0 & 0 & 0 & 0 & 0 \\ 0.160 & 0.130 & 0.070 & 0.260 & 0 & 0 & 0 & 0 & 0 \\ 0.200 & 0.170 & 0.100 & 0.340 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{17}$$

$$\mathbf{w} = \begin{bmatrix} 0.216292 \\ 0.258427 \\ 0.123596 \\ 0.174157 \\ 0.227528 \end{bmatrix} \tag{18}$$

The result (18) is coincide with the correct values (5). To obtain the correct values, in the two-cluster ANP, we need to multiply the scores of criteria to the weights which obtained by the proposed method.

5 Conclusion

In this paper an appropriate normalization method for alternatives in the AHP and the ANP was proposed. Proposed method was to divide the weights of the other alternatives by the weights of artificially added dummy alternative with perfect evaluation score. Based on the calculation results of the verification example, the proposed method was discussed. In the AHP, applying proposed method to the verification example, the result was coinciding with the correct result of the verification example. On the other hand, in the two-cluster ANP, the correct result was obtained by using the normalized weights from the proposed method. To obtain the accurate values, in this

case, we need to multiply the scores of criteria to the weights which obtained by the proposed method. Further research is improve the proposed normalization method for the ANP.

References

1. Belton, V., Gear, T.: On a Short-comming of Saaty's method of Analytic Hierarchies. Omega 11, 228–230 (1983)
2. Belton, V., Gear, T.: The legitimacy of rank reversal – A comment. Omega 13, 143–145 (1985)
3. Kinoshita, E., Nakanishi, M.: Proposal of new AHP model in light of dominant relationship among alternatives. Journal of the Operations Research Society of Japan 42, 180–197 (1999)
4. Kinoshita, E., Sugiura, S.: A comparison study of dominant AHP and similar dominant models. Journal of Research Institute of Meijo University 7, 115–116 (2008)
5. Nishizawa, K.: Simple AHP based on three-level evaluation and two-cluster ANP for the decision makers. Journal of Japanese Symposium on The Analytic Hierarchy Process 1, 97–104 (2007) (in Japanese)
6. Nishizawa, K.: Two-cluster ANP which consists of the evaluation sub-matrix of the alternatives and its transposed matrix instead of pair-wise comparisons among the criteria. Journal of Japanese Symposium on The Analytic Hierarchy Process 2, 59–67 (2010)
7. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
8. Saaty, T.L.: The Analytic Network Process. RWS Publications, Pittsburgh (1996)
9. Schoner, B., Wedley, W.C.: Ambiguous criteria weights in AHP: Consequences and solutions. Decision Sciences 20, 462–475 (1989)
10. Schoner, B., Wedley, W.C., Choo, E.U.: A unified approach to AHP with linking pins. European Journal of Operational Research 13, 384–392 (1993)

Performance Evaluation of SIRMs Models for Classification Problems

Tomoharu Nakashima and Hirosato Seki

Abstract. The performance of Single-Input Rule-Modules (SIRMs) models are studied for classification proglems. In the original version of SIRMs models, each fuzzy if-then rule has a single real-valued output. The final output from an SIRMs model is discretized in application to classification problems. This paper proposes an extention to the SIRMs models where each fuzzy if-then rule has multiple real-valued outputs that represent the activation level of the corresponding classes. The classification performance of both the original and the extended models are evaluated through a series of computational experiments using two-dimensional pattern classification problems.

1 Introduction

Single Input Rule Modules (SIRMs) connected fuzzy inference model proposed by Yubazaki et al.[1] consists of several groups of single-input fuzzy if-then rules. The advantage of the SIRMs models is computational efficiency. The number of involved fuzzy if-then rules is significantly less than the conventional model where each fuzzy if-then rule has as many inputs as the problem domain at hand. The SIRMs model has been applied to the automatic control of vehicles and the stability control of inverted pendulums. Seki et al.[2, 3, 4] analyzed the property of the SIRMs inference models and also proposed extended versions of SIRMs models.

In the standard formulation of SIRMs, the weight for each rule module is considered to be constant. However, it is often necessary to have variable weights for

Tomoharu Nakashima
Osaka Prefecture University, Gakuen-cho 1-1, Naka-ku, Sakai, Osaka 599-8531, Japan
e-mail: nakashi@cs.osakafu-.ac.jp

Hirosato Seki
Osaka Institute of Technology, Omiya 5-16-1, Asahi Ward, Osaka City,
Osaka 535-8585, Japan
e-mail: h_seki@ieee.org

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 263–271.
springerlink.com

connecting rule modules in order to accommodate high-dimensional problem domains, which is the case in most of the practical cases. From this perspective, Yu et al.[5] proposed a new SIRMs model that introduces dynamic weights for rule modules and showed the effectiveness of the model for control domains.

In this paper, SIRMs connected fuzzy inference models are applied to pattern a classification problem. Two different types of SIRMs fuzzy inference models are used for this purpose. In one fuzzy inference model, each fuzzy if-then rule has only a single output value while the other fuzzy inference model employs multiple output values. The learning algorithm for both models is also derived using the gradient decent method. In the computational experiments of this paper, the performance of the proposed model is compared for a pattern classification problem.

2 Pattern Classification Problems

The standard type of pattern classification is explained in this subsection. Let us assume that a set of training patterns is given before constructing a classifier. A training pattern consists of a real-valued input vector and its corresponding target class. Consider, for example, an n-dimensional C-class pattern classification problem. It is also assumed that m training patterns $\mathbf{x}_p = (x_{p1}, x_{p2}, \ldots, x_{pn})$, $p = 1, 2, \ldots, m$ are given a priori. The task then is to construct a classifier that correctly classify an unseen pattern using the given set of the training patterns.

3 An SIRMs Fuzzy Inference Model with a Single Output

In this section, an SIRMs fuzzy inference model with a single output (SIRMs with a single output) is presented. This model is the first version of SIRMs variants. It is assumed that each SIRM considers only one attribute in the antecedent part of fuzzy if-then rules for a given problem. Thus, for an n-dimensional pattern classification problem, there are n SIRMs in an entire fuzzy model. Fuzzy if-then rules of the SIRMs fuzzy inference model are as follows:

$$\text{Rules-}i\text{: } \{x_i = A_i^j \longrightarrow y = y_i^j\}_{j=1}^{m_i}, \quad i = 1, \ldots, n, \tag{1}$$

where Rules-i stands for the i-th single-input rule module that contains m_i fuzzy if-then rules, x_i corresponds to the i-th attribute of an input vector, A_j^i is the antecedent fuzzy set of the j-th rule in Rules-i, and y_i^j is the consequent real value of the j-th fuzzy if-then rule in Rules-i.

Let us assume that an input vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given to an SIRMs fuzzy inference model. Then, the i-th attribute x_i of the input vector is used to calculate the output value y_i in Rules-i as follows:

$$y_i = \frac{\sum_{j=1}^{m_i} A_i^j(x_i) \cdot y_i^j}{\sum_{j=1}^{m_i} A_i^j(x_i)}, \tag{2}$$

where $A_i^j(\cdot)$ is the membership function that calculates the compatibility degree of the antecedent part A_i^j in the j-th fuzzy if-then rule in Rules-i.

Final inference result y of the proposed SIRMs model is given by

$$y = \sum_{i=1}^{n} w_i \cdot y_i, \tag{3}$$

where w_i is the weight assigned to Rules-i. The weight represents the importance degree of attributes. Thus, if the value of w_i is large, the influence of the i-th attribute of input value will be high.

3.1 Learning Algorithm

In this section the learning algorithm for the SIRMs fuzzy inference model is described. The learning algorithm is derived based on the gradient decent concept where the modification rules for the parameters of the SIRMs fuzzy inference model is constructed so that a pre-defined error is minimized. In this paper the error function is defined as follows:

$$\varepsilon = \frac{1}{2}(y^T - y)^2, \tag{4}$$

where y^T is the target output value and y is the actual output from the SIRMs model whose parameters are to be tuned so that the above error function is minimized.

Let us assume that we use a triangular-type membership function for the antecedent part of fuzzy if-then rules as follows:

$$A_i^j(x_i) = \begin{cases} 1 - \dfrac{|x_i - a_i^j|}{b_i^j}, & a_i^j - b_i^j \le x_i \le a_i^j + b_i^j, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where a_i^j and b_i^j are the center and the width of the fuzzy set $A_i^j(x_i)$, respectively, and x_i is an input value for the i-th attribute. The update rules for a_i^j, b_i^j, and y_i^j at the $(t+1)$-th iteration are defined as follows:

$$w_i^{new} = w_i^{old} + \alpha \cdot (y^T - y) \cdot y_i, \tag{6}$$

$$a_i^{j,new} = a_i^{j,old} + \beta \cdot (y^T - y) \cdot w_i^{old} \cdot \frac{y_i^j - y_i^{old}}{\sum_{j=1}^{m_i} A_i^{j,old}(x_i)} \cdot \frac{\mathrm{sgn}(x_i - a_i^{j,old})}{b_i^{j,old}}, \tag{7}$$

$$b_i^{j,new} = b_i^{j,old} + \gamma \cdot (y^T - y) \cdot w_i^{old} \cdot \frac{y_i^j - y_i^{old}}{\sum_{j=1}^{m_i} A_i^{j,old}(x_i)} \cdot \frac{|x_i - a_i^{j,old}|}{(b_i^{j,old})^2}, \tag{8}$$

$$y_i^{j,new} = y_i^{j,old} + \eta \cdot (y^T - y) \cdot w_i^{old} \cdot \frac{A_i^{j,old}(x_i)}{\sum_{j=1}^{m_i} A_i^{j,old}}, \tag{9}$$

where y^T is the target output value, α, β, γ are learning rates, t represents the number of learning iterations, and sgn is a signum function that is defined as follows:

$$\mathrm{sgn} = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -1, & \text{otherwise.} \end{cases} \tag{10}$$

4 SIRMs Fuzzy Inference Models with Multiple Consequent Values

The SIRMs fuzzy inference model described in the previous sections is its initial version among a number of variants. Each fuzzy if-then rule of each rule module has only one attribute in the antecedent part and only one real output in the consequent part. This section presents an extention to the SIRMs fuzzy inference model so that multiple real values are involved in fuzzy if-then rules. Now the rule modules in the extended SIRMs fuzzy inference model can be written as follows for an n-dimensional C-class pattern classification problem:

$$\text{Rules-}i\text{: } \left\{ x_i = A_i^j \longrightarrow \begin{pmatrix} y^1 \\ y^2 \\ \cdots \\ y^C \end{pmatrix} = \begin{pmatrix} y_i^{j,1} \\ y_i^{j,2} \\ \cdots \\ y_i^{j,C} \end{pmatrix} \right\}_{j=1}^{m_i}, \ i = 1, 2, \cdots, n. \tag{11}$$

In the consequent part of fuzzy if-then rules in the extended SIRMs fuzzy inference model, there are C real values which represent the activation level of the corresponding class for an input vector. This idea is also used in the multi-layered perceptron. Thus the number of parameters for each fuzzy if-then rules is C times larger than that of the initial SIRMs fuzzy inference model.

The output from the extended SIRMs model is C-dimensional real-vector. The dimensionality of the output vector can be determined as the predefined number of classes for the pattern classification problem in hand. Each element of the output

vector is calculated in the same way as in the original SIRMs fuzzy inference model as follows:

$$\mathbf{y} = (y_1, y_2, \ldots, y_C), \tag{12}$$

where

$$y_k = \sum_{i=1}^{n} w_i^k \cdot y_i^k, \quad k = 1, 2, \ldots, C, \tag{13}$$

where w_i^k is the k-th weight assigned to the i-th SIRM, and y_i^k is the k-th output from the i-th SIRM.

The learning algorithm for the extended SIRMs fuzzy inference model is basically same as that of the initial one. That is, the error value shold be minimized by modifying the parameters that are involved in the extended SIRMs fuzzy inference model. Since the difference between the initial SIRMs model and the extended one is only in the consequent part, Equations (7) and (8) are the same. The only thing is to change the update rule (2) to accommodate multiple values. The modified update rule is written in the following:

$$y_i^{j,k,new} = y_i^{j,k,old} + \gamma \cdot (y^{T,k} - y^k) \cdot w_i^{old} \cdot \frac{A_i^{j,old}(x_i)}{\sum_{j=1}^{m_i} A_i^{j,old}}, \quad k = 1, 2, \cdots, C. \tag{14}$$

5 Computational Experiments

This section examines the performance of SIRMs models for pattern classification. An OR problem is used in this section for this purpose. The OR problem contains four training patterns from two classes (Class 1 and Class 2). The problem is shown in Fig. 1. There are four two-dimensional training patterns from two classes. The four training patterns takes either 0.0 or 1.0 for each attribute. The training pattern that is situated in the origin (i.e., $(0,0)$) comes from Class 1 while the other three patterns are from Class 2.

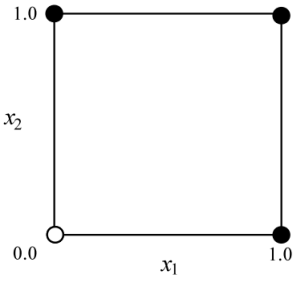


Fig. 1 An OR classification problem.

In the learning of SIRMs fuzzy inference models, the information on the training patterns should be decoded so that they are accommodated in their learning frameworks. For the SIRMs fuzzy inference model with a single real-valued consequent value, the target output for the Class 1 pattern is transferred into 0.0 while that for Class 2 patterns are specified as 1.0. After the learning algorithm of SIRMs fuzzy inference models, the output value for an unseen input value is decoded to Class 1 if the output value is less than or equal to 0.5, and it is decoded into Class 2 otherwise. For the learning of SIRMs fuzzy inference models with multiple consequent values, the number of consequent values is specified as two (i.e., the same number as the number of classes of the problem). The target vector for the Class 1 training pattern is specified as $(1.0, 0.0)$ and that for the Class 2 training patterns as $(0.0, 1.0)$. In determining the class of an unseen input pattern by this model after learning, the class with the larger output value is used as the inference results. We specified the experimental settings as follows: the weight of each SIRM is set to 1.0, each consequent value is set to 0.0, the learning rates are set to $(\alpha, \beta, \gamma, \eta) = (0.1, 0.1, 0.01, 0.01)$, and the Gaussian and triangular functions are used.

Figure 2 shows the experimental results for the SIRMs fuzzy inference model with a single consequent. The figure shows the error between the actual output from the SIRMs fuzzy inference model and the target value during the course of the learning. The error is defined as the sum of the difference between the target output and actual output for the output values from SIRMs inference models. Although different number of fuzzy partitions for each axis were used in the experiments, there was no difference as the training patterns are only placed in the extreme values of each

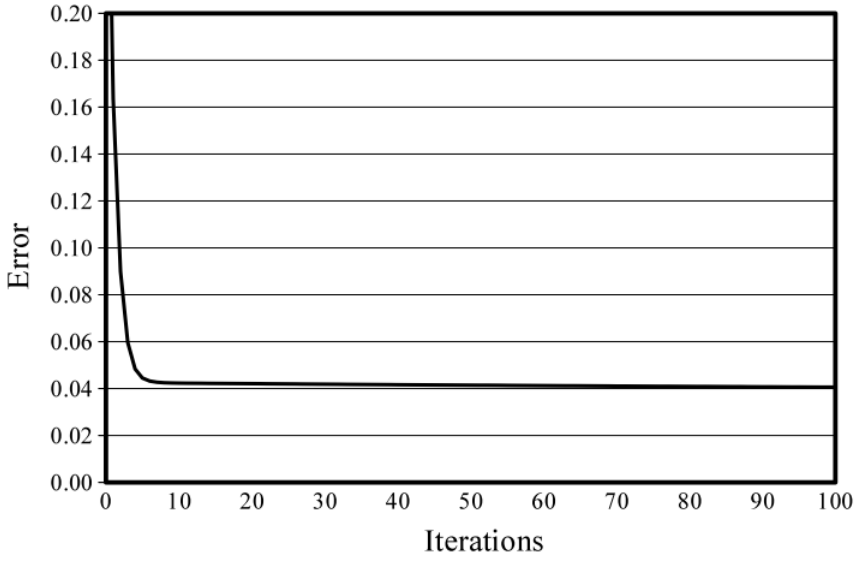


Fig. 2 SIRMs with a single consequent value (Triangular fuzzy sets).

axis (i.e., 0.0 and 1.0). Thus those fuzzy sets that do not cover any training patterns are not modified during the learning. From this figure we can see that the error value decreases as the learning proceeds, however it saturates at around 0.04.

Figure 3 shows the experimental results where the same configurations as in Fig. 2 except Gaussian-type fuzzy sets instead of triangular-type ones. The classification results with different numbers of fuzzy partitions (i.e., 2, 3, 4, and 5) are used in the computational experiments. Different results were obtained for different number of fuzzy partitions as any Gaussian-type fuzzy sets covers any training patterns even with a small amount of compatibility degrees. However, it can be seen that the error values saturates at the similar values to the case in Fig. 2. Thus the type of the fuzzy sets that are used in SIRMs model does not have a strong influence on the classification performance for the OR problem.

Figure 4 shows the experimental results for the SIRMs fuzzy inference model with multiple consequent values. The number of consequent values for each SIRM is specified as two for this two-dimensional problem. From this figure, it can be noted that the error values is about the twice as the cases in a single output consequent value (Figs. 2 and 3). This is because the error value was summed for all output values from the SIRMs fuzzy inference model. In this case, since the number of output values is twice as that of the case in a single output values, the error values are also almost doubled. It should be also mentioned that the triangular-type functions with only the two extreme positions were modified during the learning.

The experimental results for the SIRMs fuzzy inference models with multiple consequent values and Gaussian-type fuzzy sets are shown in Fig. 5. The same

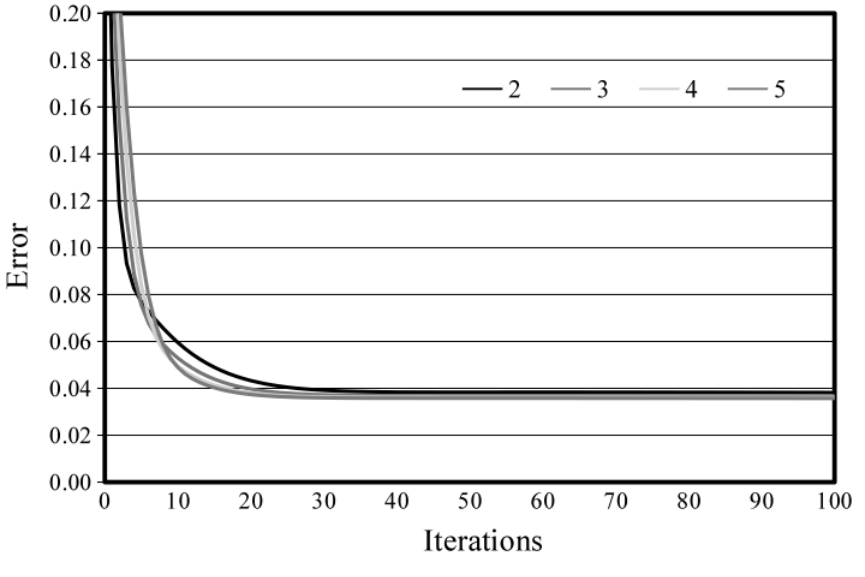


Fig. 3 SIRMs with a single consequent value (Gaussian fuzzy sets).

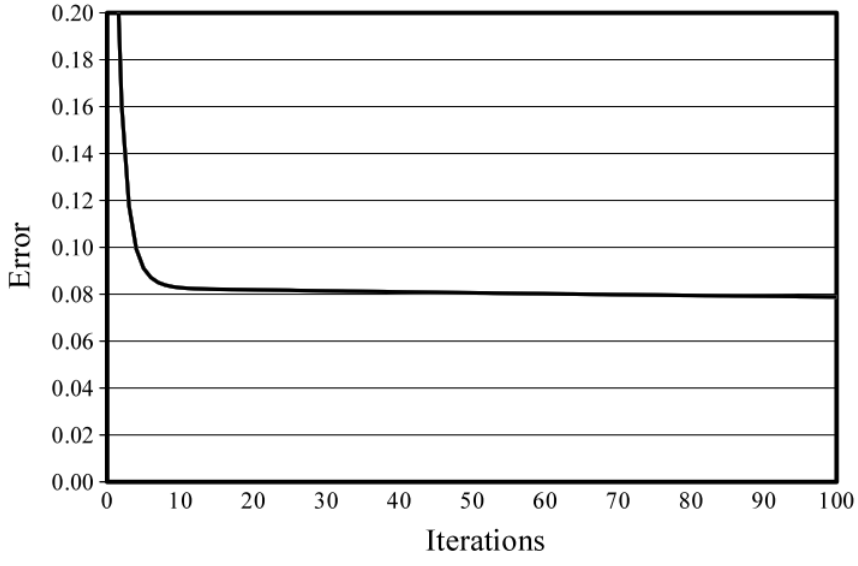


Fig. 4 SIRMs with multiple consequent values (Triangular fuzzy sets).

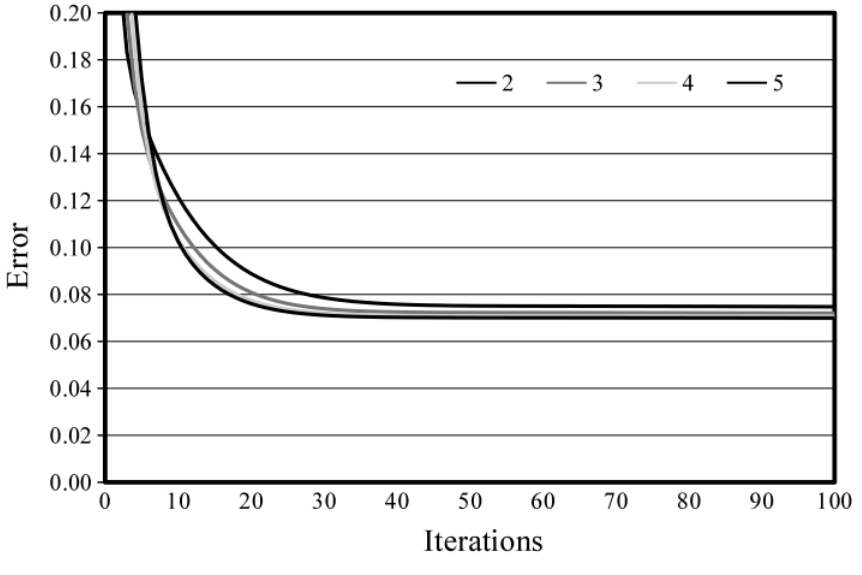


Fig. 5 SIRMs with multiple consequent values (Gaussian fuzzy sets).

discussion can be made as in the case of the other configurations. That is, the classification performance is not under the influence of the type of fuzzy sets and the number of consequent values in SIRMs for the OR problem.

6 Conclusions

In the conventional SIRMs model, the final inference result of the model for an input pattern is calculated as the weighted sum of the inference result from each rule module. Since the weight for each rule module is constant, the SIRMs model cannot be applied to highly dynamic problem domains. The learning algorithm for the proposed SIRMs model is also derived using gradient decent framework. The learning algorithm enables to automatically construct an appropriate input-output mapping by the proposed SIRMs model from a set of training patterns.

Future works include other formulation such as linear polynomials should be evaluated.

References

1. Yubazaki, N.: Sirms dynamically connected fuzzy inference model and its applications. In: Proc. IFSA 1997, Prague, Czech, vol. 3, pp. 410–415 (1997), http://ci.nii.ac.jp/naid/10018385046/en/
2. Seki, H., Ishii, H., Mizumoto, M.: On the generalization of single input rule modules connected type fuzzy reasoning method. IEEE Transactions on Fuzzy Systems 16(5), 1180–1187 (2008), doi:10.1109/TFUZZ.2008.924182
3. Seki, H., Ishii, H., Mizumoto, M.: On the monotonicity of fuzzy-inference methods related to t-s inference method. IEEE Transactions on Fuzzy Systems 18(3), 629–634 (2010), doi:10.1109/TFUZZ.2010.2046668
4. Seki, H., Mizumoto, M.: On the equivalence conditions of fuzzy inference methods — part 1: Basic concept and definition. IEEE Transactions on Fuzzy Systems 19(6), 1097–1106 (2011), doi:10.1109/TFUZZ.2011.2160268
5. Yi, J., Yubazaki, N., Hirota, K.: A proposal of sirms dynamically connected fuzzy inference model for plural input fuzzy control. Fuzzy Sets and Systems 125(1), 79–92 (2002), http://www.sciencedirect.com/science/article/pii/S0165011400001354, doi:10.1016/S0165-0114(00)00135-4

Ranking Alternatives under FWA by Riemann Integral Based Mean of Removals

Ta-Chung Chu[1,*] and Sophy Lach[2]

Abstract. This paper suggests developing membership functions of the final fuzzy evaluation values of alternatives under fuzzy weighted average (FWA), where ratings of alternatives versus qualitative criteria and the importance weights of all criteria are assessed in linguistic values represented by fuzzy numbers. Ranking approach of Riemann integral based mean of removals is applied to solve defuzzification to complete the model. Solving square root of affected quadratic equation through Riemann integral in using mean of removals for obtaining the integral values under membership functions from FWA will be studied. Formulae of ranking approach can be clearly developed, making computation execution of the proposed model more efficient. A numerical example has demonstrated feasibility of the proposed model.

Keywords: Ranking, FWA, Riemann integral, Mean of removals.

1 Introduction

Fuzzy MCDM has been widely applied to resolve many problems under uncertain environment. A review of fuzzy MCDM can be seen in [4,7,14]. Fuzzy weighted average (FWA) is one of the most effective techniques in fuzzy MCDM. Many works have been studied in FWA. In 1987, Dong and Wang [5] considered the

Ta-Chung Chu
Graduate Institute of Industrial Management, Department of Business Administration, Southern Taiwan University, Taiwan
e-mail: tcchu@mail.stut.edu.tw

Sophy Lach
Department of Business Administration, Southern Taiwan University, Taiwan

* Corresponding author.

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 273–286.
springerlink.com

computational aspect of sup-min convolution when applied to weighted average operations. They used a computational algorithm base on α-cut representation of fuzzy set, a nonlinear programming implementation of the extension principle, and interval analysis. Their method provides a precise but distinct solution to the FWA. Their method was further improved by subsequent works in [8-10, 12-13].

Despite the merits, the above FWA methods cannot rank similar fuzzy numbers efficiently due to what they produced are approximate membership functions. Their calculations will be too tedious if many qualitative and quantitative criteria must be considered. These limitations certainly hinder their applicability to real world problems. The degree of accuracy in approximation depends on the discretization refinement of the α-cut of fuzzy numbers. To resolve the above problems, this research suggests a FWA model, where ratings of alternatives versus qualitative criteria and the importance weights of all criteria are assessed in linguistic values represented by fuzzy numbers. Quantitative criteria are categorized to benefit and cost. The membership functions of the final fuzzy evaluation values from FWA can be developed through α-cut and interval arithmetic. The final evaluation values from the proposed model are still fuzzy numbers. Thus a ranking approach is needed.

A comparison and review of ranking methods can be seen in [1-2]. Herein the intuitive method, mean of removals [11], is used to defuzzify all the final fuzzy numbers. Riemann integral will be embedded to mean of removals, which can help obtain the areas under the left and right membership functions in order to produce defuzzified values to complete the model. Defuzzification formulas can be clearly developed, which makes the execution of the proposed model more efficient and the applicability of fuzzy decision making more expandable.

2 Fuzzy Set Theory

Zadeh (1965) [17] initiated the use of fuzzy set theory to deal with problem involving fuzzy phenomena which conveys imprecision and subjectivity of the human decision making. In a universe of discourse *X*, a fuzzy subset *A* is defined with a membership function $f_A(x)$ that maps each element x in *X* to a real number in the interval [0, 1]. A fuzzy subset *A* in *X* is a set of ordered pairs which can be expressed as $A = \{(x, f_A(x))/\ x \in X\}$ where *X* is the universe of discourse, x is an element in *X*, *A* is a fuzzy subset in *X*, $f_A(x)$ is the membership function of x in *A* [11].

2.1 Fuzzy Number

A real fuzzy number *A* is described as any fuzzy subset of the real line *R* with membership function f_A which processes the following properties [6]:

(a) f_A is a continuous mapping from R to the closed interval [0,1]

(b) $f_A(x)=0,$ for all $x\in(-\infty,a]$;

(c) f_A is strictly increasing on [a, b];

(d) $f_A(x)=1,$ for all $x\in[b,c]$;

(e) f_A is strictly decreasing on [c, d];

(f) $f_A(x)=0,$ for all $x\in(d,\infty]$

where $a,b,c,d\in R$. It is assumed that A is convex, normal and bounded. A general/trapezoidal fuzzy number can be denoted as $A=[a,b,c,d]/A=(a,b,c,d)$.

2.2 α-cut and Interval Arithmetic

The α-cut of a fuzzy number A can be defined as $A^\alpha=\{x|f_A(x)\geq\alpha\}$, where $x\in R$, $\alpha=[0,1]$. A^α is a non-empty bounded closed interval contained in R and can be denoted by $A^\alpha=[A_l^\alpha,A_u^\alpha]$. Given fuzzy numbers A and B, A, B, $r\in R^+$, via interval arithmetic, some main operations can be expressed as follows [11]:

$$(A\oplus B)^\alpha=[A_l^\alpha+B_l^\alpha,A_u^\alpha+B_u^\alpha] \tag{1}$$

$$(A\ominus B)=[A_l^\alpha-B_u^\alpha,A_u^\alpha-B_l^\alpha] \tag{2}$$

$$(A\otimes B)^\alpha=[A_l^\alpha.B_l^\alpha,A_u^\alpha.B_u^\alpha] \tag{3}$$

$$(A\oslash B)^\alpha=\left[\frac{A_l^\alpha}{B_u^\alpha},\frac{A_u^\alpha}{B_l^\alpha}\right] \tag{4}$$

2.3 Linguistic Values and Fuzzy Numbers

Linguistic values are those values represented in words or sentences in natural or artificial languages, where each linguistic value can be modeled by a fuzzy set [16]. In fuzzy set theory, conversion scales are employed to convert linguistic values into fuzzy numbers. Determining number of conversion scales is generally intuitive and subjective. Linguistic values can be represented by trapezoidal fuzzy numbers by decision-makers based on system needs or experience for arithmetic operations. For instance, the linguistic value "importance" can have linguistic values and trapezoidal fuzzy numbers with quadruplets such as Extremely Importance (EI)=(0.7,0.9,1,1), Very Important (VI)=(0.5,0.7,0.8,1), Important (I)=(0.3,0.45,0.55,0.7), Fairly Important (FI)=(0,0.2,0.3,0.5), Slightly Important (SI)=(0,0,0.1,0.3).

2.4 Defuzzification

Mean of removals [11] is used to complete the proposed method. Consider a general fuzzy number A=[a, b, c, d], $A \in R^+$. The left removal, denoted by A_L, the right removal, denoted by A_R, and the mean of removals of A, are defined as the following equations. $M(A)$ is used to rank fuzzy numbers. Larger $M(A)$ leads to larger A.

$$A_L = b - \int_a^b f_A^L(x)dx, \quad A_R = c + \int_c^d f_A^R(x)dx \tag{5}$$

$$M(A) = \frac{1}{2}(A_L + A_R) = \frac{1}{2}(b+c) + \frac{1}{2}\left(\int_c^d f_A^R(x)dx - \int_a^b f_A^L(x)dx\right) \tag{6}$$

3 Model Development

3.1 Determining the Alternatives and Criteria

Assume that a committee of k decision makers (*i.e.* D_t, t=1~k) is formed to select m alternatives (*i.e.* A_i, i=1~m) and determine n selection criteria (*i.e.* C_j, j=1~n). Moreover, assume that the importance weights of the criteria and the performance ratings under each qualitative criterion are assessed in linguistic values represented by positive trapezoidal fuzzy numbers.

3.2 Aggregating the Importance Weights

Let $w_{jt} = (a_{jt}, b_{jt}, c_{jt}, d_{jt}), w_{jt} \in R^+, \quad j = 1,2,...,n, \quad t = 1,2,...,k,$ be the importance weight given by decision-maker D_t to criterion C_j. The aggregated importance weight, w_j of criterion C_j assessed by the committee of k decision-makers can be evaluated as:

$$w_j = \left(\frac{1}{k}\right) \otimes \left(w_{j1} \oplus w_{j2} \oplus ... \oplus w_{jk}\right),$$

$$a_j = \sum_{t=1}^{k} \frac{a_{jt}}{k},\ b_j = \sum_{t=1}^{k} \frac{b_{jt}}{k},\ c_j = \sum_{t=1}^{k} \frac{c_{jt}}{k},\ d_j = \sum_{t=1}^{k} \frac{d_{jt}}{k}. \tag{7}$$

3.3 Aggregating the Ratings of Alternatives versus Qualitative Criteria

Let $x_{ijt} = (o_{ijt}, p_{ijt}, q_{ijt}, r_{ijt}), x_{ijt} \in R^+, \quad i = 1,2,...,m, \quad j = 1,2,...,n,$ $t = 1,2,...,k,$ be the linguistic rating assigned to A_i by D_t under C_j. The

aggregated ratings, x_{ij} of A_i under C_j assessed by k decision-makers can be evaluated as:

$$x_{ij} = \left(\frac{1}{k}\right) \otimes \left(x_{ij1} \oplus x_{ij2} \oplus \ldots \oplus x_{ijk}\right),$$

$$o_{ij} = \sum_{t=1}^{k} \frac{o_{ijt}}{k},\ p_{ij} = \sum_{t=1}^{k} \frac{p_{ijt}}{k},\ q_{ij} = \sum_{t=1}^{k} \frac{q_{ijt}}{k},\ r_{ij} = \sum_{t=1}^{k} \frac{r_{ijt}}{k}. \tag{8}$$

Linguistic values and fuzzy numbers such as Very Low (VL)=(0,0,0.1,0.3), Low (L)=(0,0.2,0.3,0.5), Medium (M)=(0.3,0.45,0.55,0.7), High (H)=(0.5,0.7,0.8,1), Very High (VH)=(0.7,0.9,1,1) can be used. In a special case, for qualitative cost criteria (the lower, the better), such as risk, reverse linguistic values and fuzzy numbers, such as Very Low (VL)=(0.7,0.9,1,1), Low (L)= (0.5,0.7,0.8,1), Medium (M)=(0.3,0.45,0.55,0.7), High (H)=(0,0.2,0.3,0.5), Very High (VH)=(0,0,0.1,0.3), are needed to maintain consistency to the whole assessment of other criteria.

3.4 Rating of Alternatives versus Quantitative Criteria

Quantitative criteria values r_{ij} may have different units. Thus all ratings of alternatives versus quantitative criteria must be normalized into a comparable scale for calculation rationale. Herein Chen's method [3] is applied as follows:

$$x_{ij} = \left(\frac{a_{ij}}{d_j^*}, \frac{b_{ij}}{d_j^*}, \frac{c_{ij}}{d_j^*}, \frac{d_{ij}}{d_j^*}\right), \quad j \in B \ , \ x_{ij} = \left(\frac{a_j^-}{d_{ij}}, \frac{a_j^-}{c_{ij}}, \frac{a_j^-}{b_{ij}}, \frac{a_j^-}{a_{ij}}\right), \quad j \in C \tag{9}$$

where $d_j^* = \max_i d_{ij}$ if $j \in B$, $a_j^- = \min_i a_{ij}$ if $j \in C$, x_{ij} denotes the normalized value of r_{ij} , B and C denotes the benefit and cost criteria, respectively.

3.5 Developing Membership Function for FWA

A fuzzy weighted average (FWA) is shown as follows:

$$\bar{y}_i = \frac{\left(w_1 \otimes x_{i1}\right) \oplus \ldots \oplus \left(w_j \otimes x_{ij}\right) \oplus \ldots \oplus \left(w_n \otimes x_{in}\right)}{w_1 \oplus \ldots \oplus w_j \oplus \ldots \oplus w_n} \tag{10}$$

Via Eqs. (1)~(4), the membership functions for FWA (*i.e.* $\bar{y}_i$ in (10)) can be developed. Firstly, $\bar{y}_i^{\alpha} = \left[\bar{y}_{il}^{\alpha}, \bar{y}_{iu}^{\alpha}\right]$ and

$$\bar{y}_{il}^{\alpha} = \frac{\left(\sum_{j=1}^{n}(b_j - a_j)(p_{ij} - o_{ij})\right)\alpha^2 + \left(\sum_{j=1}^{n}\left[a_j(p_{ij} - o_{ij}) + o_{ij}(b_j - a_j)\right]\right)\alpha + \sum_{j=1}^{n} a_j o_{ij}}{\sum_{j=1}^{n}(c_j - d_j)\alpha + \sum_{j=1}^{n} d_j} \tag{11}$$

$$\bar{y}_{iu}^{\alpha} = \frac{\left(\sum_{j=1}^{n}(c_j - d_j)(q_{ij} - r_{ij})\right)\alpha^2 + \left(\sum_{j=1}^{n}\left[d_j(q_{ij} - r_{ij}) + r_{ij}(c_j - d_j)\right]\right)\alpha + \sum_{j=1}^{n} d_j r_{ij}}{\sum_{j=1}^{n}(b_j - a_j)\alpha + \sum_{j=1}^{n} a_j} \tag{12}$$

Let $\sum_{j=1}^{n} a_j o_{ij} = U_i$, $\sum_{j=1}^{n} b_j p_{ij} = V_i$, $\sum_{j=1}^{n} c_j q_{ij} = Y_i$, $\sum_{j=1}^{n} d_j r_{ij} = Z_i$

$\sum_{j=1}^{n} a_j = H_1$, $\sum_{j=1}^{n} b_j = H_2$, $\sum_{j=1}^{n} c_j = H_3$, $\sum_{j=1}^{n} d_j = H_4$, $\sum_{j=1}^{n}(b_j - a_j) = G_1$, $\sum_{j=1}^{n}(c_j - d_j) = G_2$

$\sum_{j=1}^{n}(b_j - a_j)(p_{ij} - o_{ij}) = E_{i1}$, $\sum_{j=1}^{n}\left[a_j(p_{ij} - o_{ij}) + o_{ij}(b_j - a_j)\right] = F_{i1}$

$\sum_{j=1}^{n}(c_j - d_j)(q_{ij} - r_{ij}) = E_{i2}$, $\sum_{j=1}^{n}\left[d_j(q_{ij} - r_{ij}) + r_{ij}(c_j - d_j)\right] = F_{i2}$

Eqs. (11) and (12) lead to the following two equations to solve:

$$E_{i1}\alpha^2 + (F_{i1} - G_1 x)\alpha + U_i - H_4 x = 0 \tag{13}$$

$$E_{i2}\alpha^2 + (F_{i2} - G_2 x)\alpha + U_i - H_1 x = 0 \tag{14}$$

Only roots in [0,1] will be retained in (13) and (14). The left membership function, *i.e.* $f_{\bar{y}_i}^{L}(x)$, and the right membership function, *i.e.* $f_{\bar{y}_i}^{R}(x)$, of $\bar{y}_i$ can be developed as:

$$f_{\bar{y}_i}^{L}(x) = \frac{G_2 x - F_{i1} + \left[G_2^2 x^2 + I_{i1} x + J_{i1}\right]^{1/2}}{2E_{i1}}, \qquad U_i / H_4 \le x \le V_i / H_3,$$

$$1, \qquad V_i / H_3 \le x \le Y_i / H_2,$$

$$f_{\bar{y}_i}^{R}(x) = \frac{G_1 x - F_{i2} + \left[G_1^2 x^2 + I_{i2} x + J_{i2}\right]^{1/2}}{2E_{i2}}, \qquad Y_i / H_2 \le x \le Z_i / H_1. \tag{15}$$

where $I_{i1} = 4E_{i1}H_4 - 2F_{i1}G_2$, $J_{i1} = F_{i1}^2 - 4E_{i1}U_i$, $I_{i2} = 4E_{i2}H_1 - 2F_{i2}G_1$, $J_{i1} = F_{i1}^2 - 4E_{i1}U_i$, $E_{i1} \neq 0$ and $E_{i2} \neq 0$. $\bar{y}_i$ can be expressed as:

$$\bar{y}_i = \left(\frac{U_i}{H_4}, \frac{V_i}{H_3}, \frac{Y_i}{H_2}, \frac{Z_i}{H_1}; \; E_{i1}, F_{i1}, G_2, I_{i1}, J_{i1}; \; E_{i2}, F_{i2}, G_1, I_{i2}, J_{i2} \right), \; i = 1,2,...,n \tag{16}$$

3.6 Obtain Ranking Values

By Eqs. (5)~(6), the ranking value of $\bar{y}_i$ can be obtained as:

$$M(\bar{y}_i) = \frac{1}{2}\left(\frac{V_i}{H_3} + \frac{Y_i}{H_2} \right) + \frac{1}{2}\left(\int_{Y_i/H_2}^{Z_i/H_1} f_{\bar{y}_i}^R(x)dx - \int_{U_i/H_4}^{V_i/H_3} f_{\bar{y}_i}^L(x)dx \right) \tag{17}$$

where $f_{\bar{y}_i}^L(x)$ and $f_{\bar{y}_i}^R(x)$ are the left and right membership functions of $\bar{y}_i$.

Replacing left and right membership functions of Eq. (15) into (17), we have:

$$M(\bar{y}_i) = \frac{1}{2}\left(\frac{V_i}{H_3} + \frac{Y_i}{H_2} \right) + \frac{1}{2}\left(\int_{Y_i/H_2}^{Z_i/H_1} \left(\frac{G_1 x - F_{i2} + \left[G_1^2 x^2 + I_{i2} x + J_{i2} \right]^{1/2}}{2E_{i2}} \right) dx - \right.$$

$$\left. - \int_{U_i/H_4}^{V_i/H_3} \left(\frac{G_2 x - F_{i1} + \left[G_2^2 x^2 + I_{i1} x + J_{i1} \right]^{1/2}}{2E_{i1}} \right) dx \right) \tag{18}$$

First, we will solve the integral for right membership function:

$$\int_{Y_i/H_2}^{Z_i/H_1} \left(\frac{G_1 x - F_{i2} + \left[G_1^2 x^2 + I_{i2} x + J_{i2} \right]^{1/2}}{2E_{i2}} \right) dx$$

$$= \frac{G_1}{4E_{i2}} x^2 \Big|_{Y_i/H_2}^{Z_i/H_1} - \frac{F_{i2}}{2E_{i2}} x \Big|_{Y_i/H_2}^{Z_i/H_1} + \frac{1}{2E_{i2}} \int_{Y_i/H_2}^{Z_i/H_1} \left[G_1^2 x^2 + I_{i2} x + J_{i2} \right]^{1/2} dx \tag{19}$$

Assume $G_1=a$, $I_{i2}=b$, and $J_{i2}=c$, then the remaining integral part of Eq. (19) is:

$\int_{Y_i/H_2}^{Z_i/H_1} \sqrt{ax^2+bx+c}\,dx$ or in general form, we can write:

$$\int \sqrt{ax^2+bx+c}\,dx \tag{20}$$

Assume that Eq. (20) is equal to N, then Multiplying Eq. (20) with $\frac{\sqrt{ax^2+bx+c}}{\sqrt{ax^2+bx+c}}$, we have $N=\int \frac{ax^2+bx+c}{\sqrt{ax^2+bx+c}}dx$ (21)

Again, multiplying Eq. (21) with $\frac{8a}{8a}$, then we have:

$$N=\int \frac{8a^2x^2+8abx+8ac}{8a\sqrt{ax^2+bx+c}}dx$$

$$=\int\left[\left(\frac{1}{2}\sqrt{ax^2+bx+c}+\frac{x}{2}\frac{1}{2}\frac{(2ax+b)}{\sqrt{ax^2+bx+c}}\right)+\left(\frac{b}{4a}\frac{1}{2}\frac{(2ax+b)}{\sqrt{ax^2+bx+c}}\right)+\left(\frac{4ac-b^2}{8a\sqrt{ax^2+bx+c}}\right)\right]dx$$

$\int (A'B+AB')dx=\int dAB=AB$ and $\int B'dx=\int dB=B$ and in the above equation

$$A=\frac{x}{2};\quad A'=\frac{1}{2};\quad B=\sqrt{ax^2+bx+c};\quad B'=\frac{1}{2}\frac{(2ax+b)}{\sqrt{ax^2+bx+c}}.$$

Then we have (S.O.S Math 2008 [15]):

$$N=\left(\frac{x}{2}\sqrt{ax^2+bx+c}\right)+\left(\frac{b}{4a}\sqrt{ax^2+bx+c}\right)+\frac{4ac-b^2}{8a}\int\frac{dx}{\sqrt{ax^2+bx+c}} \tag{22}$$

Solve the remaining integral part of Eq. (22) through the following procedure:

$$\int\frac{dx}{\sqrt{ax^2+bx+c}}=\int\frac{2a\sqrt{ax^2+bx+c}+\sqrt{a}(2ax+b)}{2a\sqrt{ax^2+bx+c}+\sqrt{a}(2ax+b)}\times\frac{dx}{\sqrt{ax^2+bx+c}}$$

$$=\frac{1}{\sqrt{a}}\ln(2\sqrt{a}\sqrt{ax^2+bx+c}+2ax+b)\text{ (S.O.S Math 2008 [15])} \tag{23}$$

Then, replacing Eq. (23) into (22), we have:

$$N = \frac{(2ax+b)\sqrt{ax^2+bx+c}}{4a} + \frac{4ac-b^2}{8a} \times \frac{1}{\sqrt{a}} \ln(2\sqrt{a}\sqrt{ax^2+bx+c} + 2ax + b) \tag{24}$$

Replacing the values of a, b, and c into Eq. (24) and then replacing (24) into Eq. (19), the right membership function, Eq. (19), can be expressed as:

$$= \frac{G_1}{2E_{i2}}\left[\left(\frac{Z_i}{H_1}\right)^2 - \left(\frac{Y_i}{H_2}\right)^2\right] - \frac{F_{i2}}{2E_{i2}}\left(\frac{Z_i}{H_1} - \frac{Y_i}{H_2}\right) + \frac{1}{2E_{i2}} \times$$

$$\times \left[\left(\frac{\left(2G_1^2(Z_i/H_1) + I_{i2}\right)\sqrt{G_1^2(Z_i/H_1)^2 + I_{i2}(Z_i/H_1) + J_{i2}}}{4G_1^2} + \frac{4G_1^2 J_{i2} - I_{i2}^2}{8G_1^2} \times \right.\right.$$

$$\left.\left. \times \frac{1}{\sqrt{G_1^2}} \ln(2\sqrt{G_1^2}\sqrt{G_1^2(Z_i/H_1)^2 + I_{i2}(Z_i/H_1) + J_{i2}} + 2G_1^2(Z_i/H_1) + I_{i2}\right) - \right.$$

$$\left. - \left(\frac{\left(2G_1^2(Y_i/H_2) + I_{i2}\right)\sqrt{G_1^2(Y_i/H_2)^2 + I_{i2}(Y_i/H_2) + J_{i2}}}{4G_1^2} + \frac{4G_1^2 J_{i2} - I_{i2}^2}{8G_1^2} \times \right.\right.$$

$$\left.\left. \times \frac{1}{\sqrt{G_1^2}} \ln(2\sqrt{G_1^2}\sqrt{G_1^2(Y_i/H_2)^2 + I_{i2}(Y_i/H_2) + J_{i2}} + 2G_1^2(Y_i/H_2) + I_{i2}\right)\right] \tag{25}$$

Similarly, the integral value for left membership function in Eq. (18) can be solved as:

$$= \frac{G_2}{2E_{i1}}\left[\left(\frac{V_i}{H_3}\right)^2 - \left(\frac{U_i}{H_4}\right)^2\right] - \frac{F_{i1}}{2E_{i1}}\left(\frac{V_i}{H_3} - \frac{U_i}{H_4}\right) + \frac{1}{2E_{i1}} \times$$

$$\times \left[\left(\frac{\left(2G_2^2(V_i/H_3) + I_{i1}\right)\sqrt{G_2^2(V_i/H_3)^2 + I_{i1}(V_i/H_3) + J_{i1}}}{4G_2^2} + \frac{4G_2^2 J_{i1} - I_{i1}^2}{8G_2^2} \times \right.\right.$$

$$\times\frac{1}{\sqrt{G_2^2}}\ln(2\sqrt{G_2^2}\sqrt{G_2^2(V_i/H_3)^2+I_{i1}(V_i/H_3)+J_{i1}}+2G_2^2(V_i/H_3)+I_{i1}\Bigg)-$$

$$-\Bigg(\frac{(2G_2^2(U_i/H_4)+I_{i1})\sqrt{G_2^2(U_i/H_4)^2+I_{i1}(U_i/H_4)+J_{i1}}}{4G_2^2}+\frac{4G_2^2J_{i1}-I_{i1}^2}{8G_2^2}\times$$

$$\times\frac{1}{\sqrt{G_2^2}}\ln(2\sqrt{G_2^2}\sqrt{G_2^2(U_i/H_4)^2+I_{i1}(U_i/H_4)+J_{i1}}+2G_1^2(U_i/H_4)+I_{i1}\Bigg)\Bigg] \tag{26}$$

Replacing Eqs. (25) and (26) into (18), the mean of removals of each alternative (*i.e.* $M(\bar{y}_i)$) can be obtained. Obviously, if $M(A_1)>M(A_2)$, then $A_1 > A_2$; if $M(A_1)=M(A_2)$, then $A_1 = A_2$; and if $M(A_1)<M(A_2)$, then $A_1 < A_2$.

4 Numerical Example

Assume that a company must evaluate and rank its new product development projects, and a committee of four decision-makers (D_t, t=1,…,4) is responsible for evaluating five alternative projects (A_i, i=1,2,..,5) under eleven criteria (C_j, j=1,2,…,11). The eleven criteria are classified to benefic quantitative, such as potential sales (C_1), cost quantitative, such as development cost (C_2) and time to market (C_3), benefit qualitative, such as technological ability (C_4), manufacturing ability (C_5), relevance (C_6), reasonableness (C_7), product flexibility (C_8), customer satisfaction (C_9), cooperation (C_{10}),and cost qualitative, such as risk (C_{11}).

Suppose the linguistic weighting values in Section 2.3 are used by decision makers to evaluate the importance weights of criteria as shown in Table 1. Through Eq. (7), the average weights of the criteria can be obtained as also shown in Table 1. Suppose the three quantitative criteria such as potential sales (C_1), measured in million US$, development cost ($C_2$), measured in million US$, and time to market ($C_3$), measured in month, are shown in Table 2, and Eq. (9) produces their normalized values as also shown in Table 2.

Table 1 Importance Weights of Criteria and Average Weights

Criteria	Decision Makers				Average Weights
	D_1	D_2	D_3	D_4	
C_1	EI	EI	EI	EI	(0.7,0.9,1,1)
C_2	VI	VI	VI	VI	(0.5,0.7,0.8,1)
C_3	EI	VI	VI	VI	(0.55,0.75,0.85,1)
C_4	VI	EI	VI	VI	(0.55,0.75,0.85,1)
C_5	VI	I	VI	VI	(0.45,0.638,0.738,0.925)
C_6	VI	VI	EI	VI	(0.55,0.75,0.85,1)
C_7	VI	I	EI	VI	(0.5,0.688,0.788,0.925)
C_8	I	VI	I	I	(0.35,0.513,0.738,0.925)
C_9	VI	VI	I	VI	(0.45,0.638,0.738,0.925)
C_{10}	I	I	I	I	(0.3,0.45,0.55,0.7)
C_{11}	I	VI	VI	I	(0.4,0.575,0.675,0.85)

Table 2 Values of Quantitative Alternatives (Normalized Values)

Quantitative Criteria	Alternatives				
	A_1	A_2	A_3	A_4	A_5
C_1	15 (0.429)	35 (1)	10 (0.286)	25 (0.714)	15 (0.429)
C_2	6 (0.67)	10 (0.4)	4 (1)	12 (0.33)	8 (0.5)
C_3	6 (0.833)	8 (0.625)	5 (1)	10 (0.55)	5 (1)

By using linguistic values in Section 3.3, suppose the appropriate ratings of five alternatives versus eight qualitative criteria are determined as shown in Table 3. Via Eq. (8), the average ratings of alternatives can be obtained as also shown in Table 3. By Eqs. (10)-(16), membership function of $\overline{y}_i$, $i = 1,...,5$, can be developed as:

$\overline{y}_1 =$ (0.93,0.61,0.89,1.67; 0.28,1.86,-1.65,17.60,0.06;
0.18,-2.52,2.05,14.17,-0.06);
$\overline{y}_2 =$ (0.30,0.61,0.88,1.62; 0.27,1.80,-1.65,16.95,-0.07;
0.18,-2.34,2.05,13.35,-0.67);
$\overline{y}_3 =$ (0.30,0.60,0.88,1.62; 0.26,1.81,-1.65,16.30,0.15;
0.14,-2.22,2.05,12.08,0.09);
$\overline{y}_4 =$ (0.22,0.47,0.71,1.37; 0.24,1.43,-1.65,14.34,-0.11;
0.20,-2.25,2.05,13.44,-0.71);
$\overline{y}_5 =$ (0.28,0.57,0.84,1.6; 0.27,1.74,-1.65,16.67,-0.01;
0.21,-2.55,2.05,14.81,-0.46).

Table 3 Ratings of Alternatives and Average Ratings

Alternatives / Criteria		Decision Makers				Average Ratings
		D_1	D_2	D_3	D_4	
A_1	C_4	VH	H	H	H	(0.55,0.75,0.85,1)
	C_5	VH	H	H	H	(0.55,0.75,0.85,1)
	C_6	H	VH	VH	H	(0.6,0.8,0.9,1)
	C_7	H	H	M	H	(0.45,0.638,0.738,0.925)
	C_8	M	H	H	M	(0.40,0.575,0.675,0.85)
	C_9	H	VH	H	VH	(0.6,0.8,0.9,1)
	C_{10}	VH	H	VH	VH	(0.65,0.85,0.95,1)
	C_{11}	H	H	H	H	(0.5,0.7,0.8,1)
A_2	C_4	VH	H	M	H	(0.5,0.688,0.788,0.925)
	C_5	H	M	M	M	(0.35,0.513,0.61,0.775)
	C_6	VH	VH	H	VH	(0.65,0.85,0.95,1)
	C_7	H	H	H	H	(0.5,0.7,0.8,1)
	C_8	H	VH	H	VH	(0.6,0.8,0.9,1)
	C_9	H	VH	VH	H	(0.6,0.8,0.9,1)
	C_{10}	H	M	H	M	(0.4,0.575,0.675,0.85)
	C_{11}	M	M	H	H	(0.4,0.575,0.675,0.8)
A_3	C_4	VH	VH	H	H	(0.6,0.8,0.9,1)
	C_5	VH	VH	VH	VH	(0.7,09,1,1)
	C_6	VH	H	H	VH	(0.6,0.8,0.9,1)
	C_7	VH	VH	VH	VH	(0.7,09,1,1)
	C_8	L	L	L	VL	(0,0.15,0.25,0.45)
	C_9	M	H	M	L	(0.275,0.425,0.55,0.725)
	C_{10}	H	H	VH	VH	(0.6,0.8,0.9,1)
	C_{11}	H	H	M	M	(0.4,0.58,0.68,0.85)
A_4	C_4	H	M	M	M	(0.35,0.513,0.613,0.775)
	C_5	M	L	M	M	(0.225,0.388,0.448,0.65)
	C_6	VH	H	VH	VH	(0.65,0.85,0.95,1)
	C_7	M	L	M	L	(0.15,0.325,0.425,0.6)
	C_8	M	H	H	VH	(0.5,0.688,0.788,0.925)
	C_9	H	H	H	H	(0.5,0.7,0.8,1)
	C_{10}	M	H	M	M	(0.35,0.513,0.613,0.775)
	C_{11}	L	M	M	M	(0.23,0.388,0.488,0.65)
A_5	C_4	VH	H	H	VH	(0.6,0.8,0.9,1)
	C_5	VH	H	H	M	(0.5,0.688,0.788,0.925)
	C_6	H	H	VH	H	(0.55,0.75,0.85,1)
	C_7	H	H	H	H	(0.5,0.7,0.8,1)
	C_8	L	M	H	M	(0.275,0.45,0.55,0.725)
	C_9	H	M	H	M	(0.4,0.575,0.675,0.85)
	C_{10}	H	H	VH	H	(0.55,0.75,0.85,1)
	C_{11}	H	M	H	M	(0.4,0.575,0.675,0.85)

The final ranking values of alternatives can then be obtained through Eqs. (17)~(26) as $M(\bar{y}_1)=14.636$, $M(\bar{y}_2)=13.673$, $M(\bar{y}_3)=16.512$, $M(\bar{y}_4)=9.519$ and $M(\bar{y}_5)=12.382$. Obviously the ranking order of is $A_3 > A_1 > A_2 > A_5 > A_4$.

5 Conclusions

Membership functions of the final fuzzy evaluation values from FWA have been developed. Riemann integral based ranking method of mean of removals has been used to defuzzify all the fuzzy numbers to complete the model. Formulae of ranking procedure can be clearly displayed, making the execution of the proposed model more efficiently. A numerical example of ranking new products development projects of a company has demonstrated computational procedure of the suggested model.

The proposed method can also be applied to solve other multiple criteria management problems in an uncertain environment. However, when applied, the following problems are worth closer study: (1) Number of criteria can be adjusted upon different cases; (2) Further research may try to justify the feasibility of the proposed method in either an empirical study or a case study, applying it to a real company in real situations; (3) A comparison of the suggested method to the other similar ones might be added to show more convincing merits; (4) The outcome could be different if normalization formulas, linguistic ratings and weightings, number of decision maker etc. are different.

Acknowledgments. This work was supported in part by National Science Council, Taiwan, R.O.C., under Grant NSC 100-2410-H-218-002.

References

1. Abbasbandy, S., Hajjari, T.: A new approach for ranking of trapezoidal fuzzy numbers. Computers and Mathematics with Applications 57(3), 413–419 (2009)
2. Asady, B.: Revision of distance minimization method for ranking of fuzzy numbers. Applied Mathematical Modeling 35, 1306–1313 (2011)
3. Chen, C.T.: A fuzzy approach to select the location of the distribution center. Fuzzy Sets and Systems 118(1), 65–73 (2001)
4. Chu, T.C., Lin, Y.: An extension to fuzzy MCDM. Computers and Mathematics with Applications 57(3), 445–454 (2009)
5. Dong, W.M., Wang, F.S.: Fuzzy weighted averages and implementation of the extension principle. Fuzzy Sets and Systems 21(2), 183–199 (1987)
6. Dubois, D., Prade, H.: Operation on fuzzy number. International Journal of System Science 9(6), 613–626 (1978)
7. Dursun, M., Karsak, E.E.: A fuzzy MCDM approach for personnel selection. Expert Systems with Applications 37(6), 4324–4330 (2010)

8. Guh, Y.Y., Hon, C.C., Wang, K.M., Lee, E.S.: Fuzzy weighted average: a max-min paired elimination method. Computers and Mathematics with Applications 32(8), 115–123 (1996)
9. Guu, S.M.: The fuzzy weighted average method revisited. Fuzzy Sets and Systems 126(3), 411–414 (2002)
10. Kao, C., Liu, S.-T.: Fractional programming approach to fuzzy weighted average. Fuzzy Sets and Systems 120(3), 435–444 (2001)
11. Kaufmann, A., Gupta, M.M.: Introduction to fuzzy arithmetic-theory and application. Van Nostrand Reinhold, New York (1991)
12. Lee, D.H., Park, D.: An efficient algorithm for fuzzy weighted average. Fuzzy Sets and Systems 87(1), 39–45 (1997)
13. Liou, T.-S., Wang, M.-J.: Fuzzy weighted average: an improved algorithm. Fuzzy Sets and Systems 49(3), 307–315 (1992)
14. Momeni, M., Maleki, M.H., Afshari, M.A., Moradi, J.S., Mohammadi, J.: A fuzzy MCDM approach for evaluating listed private banks in Tehran Stock Exchange based on balanced scorecard. International Journal of Business Administration 2(1), 80–97 (2011)
15. Math, S.O.S.: Integrals containing the square root of ax2+bx+c (2008), http://www.sosmath.com/tables/integral/integ15/integ15.html
16. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Information Sciences 8(3), 199–249 (I), 301–357 (II) (1975)
17. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)

Representation and Evaluation of Granular Systems*

Marcin Szczuka and Dominik Ślęzak

Abstract. We put forward a general framework for looking at the quality of granular systems. We discuss issues of representing and using granular methods in comparison with traditional approach. We focus on the demand for quality evaluation strategies that judge granular environments by their ability to reduce computational effort and simplify description, assuring required level of precision and relevance at the same time. We provide several examples to illustrate our approach.

1 Introduction

The idea of granular systems and granular computing builds on general observation, that in many real-life situations we are unable and, actually, not willing to precisely discern between similar objects. Our perception of such

Marcin Szczuka
Institute of Mathematics, The University of Warsaw, Banacha 2,
02-097 Warsaw, Poland
e-mail: szczuka@mimuw.edu.pl

Dominik Ślęzak
Infobright Inc., Krzywickiego 34 lok. 219, 02-078 Warsaw, Poland
and Institute of Mathematics, The University of Warsaw
Banacha 2, 02-097 Warsaw, Poland
e-mail: slezak@infobright.com

* This work was supported by the grant N N516 077837 from the Ministry of Science and Higher Education of the Republic of Poland, the Polish National Science Centre grant 2011/01/B/ST6/03867 and by the Polish National Centre for Research and Development (NCBiR) under Grant No. SP/I/1/77065/10 in frame of the the strategic scientific research and experimental development program: "Interdisciplinary System for Interactive Scientific and Scientific-Technical Information".

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 287–296.
springerlink.com

universe is *granular*, which means, that we are dealing with groups of objects (*granules*) with limited resolution.

The granules come in all shapes and sizes. Granulation may be introduced into universe of discourse in great many ways, as exemplified by the diversity of approaches presented in [14]. The granules can be created as snapshots of information, such as sets and intervals, fuzzy and rough subsets, cluster and pattern descriptions, as well as various other constructs over the original universe of investigated objects [22]. They can also take a form of the sets of original or granulated objects gathered together according to some constraints, with various types of links and structures spanned over them.

The granules, whatever their origin, may provide the means to simplify, enhance, and speed up the computational tasks such as: searching, mining, or reasoning. We expect that by replacing original objects with granules we may obtain more compact outcome and avoid too detailed, hence non-transparent results, as well as – what is equally important in many applications – unnecessarily complex calculations leading to those results. This is helpful wherever the most fine-grained result descriptions are not required. In some other situations, the results computed at the granular level may be actually of better quality. Yet another motivation for using granules stems from their potential for reducing the computational effort needed to obtain the results, if more efficient granular operations result in the same or almost the same outcomes.

The use of granules comes, of course, at a cost. In granular environments, we are often faced with considerable imprecision or vagueness, which, if not handled carefully, can cause inefficiencies of computational processes, both with respect to their speed and their accuracy. The key to success is to properly manage the quality associated with each given level of granularity. In particular, one needs to take care of controlling the growth of imprecision along the chain of granular computations (see e.g. [19]) and the specifics of those computations (see e.g. [20]), as well as optimizing the quality of granular systems both at the point of their construction (see e.g. [2], [17]) and application-specific deployment (see e.g. [18]).

In this paper, we focus mainly on the last two of the above aspects. We consider the framework proposed in [13], where some fundamental quality measures, such as experimental justification and specificity (precision), have been formulated. Using a number of examples taken from the domains such as data mining [16], data processing [17], and document retrieval, we show that those measures are indeed worth considering while evaluating and optimizing the granular systems. However, we also claim that construction of a highly-scoring granular system cannot be the goal in itself.

The ultimate criterion of success should be related more directly to the tangible requirements of particular domains of applications, treating the previously-introduced measures rather as the means for heuristic estimate of the efficiency of granular solutions. In order to better grasp it, we propose a general way of looking at quality of the system. In this approach build a high level quality measure that assesses the overall performance of the

granular system in the way similar to empirical risk function known from the classical statistical learning theory (see [21]).

We begin the paper with introduction of notation and vocabulary needed top resent the main idea, i.e., the scheme for granule system evaluation in (Section 2). The main idea is illustrated with examples of diverse granular systems in Section 3. We conclude with general remarks and discussion of possibilities of future applications of the proposed scheme.

2 Granule System Quality Evaluation

We also refer to some analogies to the filter/wrapper/embedded methodologies in feature subset selection [8] and the principles of optimization of controllers in the complex control systems, where the common goal is to design a solution at an intermediate stage of a longer process.

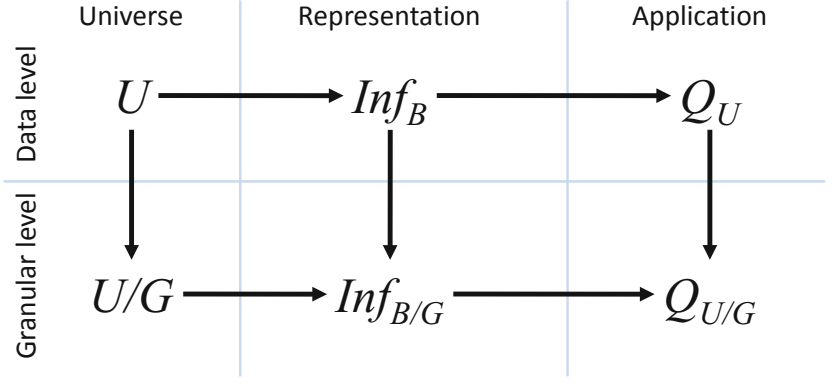


Fig. 1 The layout of proposed scheme. U is the universe containing objects by *information vectors* $Inf_B(u) = \langle a_1(u), \ldots, a_m(u) \rangle$, where $a_1, \ldots, a_n \in B$ are attributes, i.e., mappings that assign attribute values to objects.

To better focus the presentation, we use Figure 1 to introduce a little bit of notation that we will further use to characterize granular systems. By U we denote the universe of discourse. In the most typical setting U contains objects that are represented by means of their *information vectors* $Inf_B(u) = \langle a_1(u), \ldots, a_m(u) \rangle$, where $a_1, \ldots, a_n \in B$ are attributes, i.e., mappings that assign attribute values to objects. This representation follows [12], but depending on application, U may gather, e.g., tuples in a relational database, rows in a data sample, pixels of an image, measurements within a time series, et cetera.

When we start discussing the granules, we have to modify the plain information vector in such a way that its coordinates represent information about

granules of objects, not just single instances. For this purpose we introduce the granular system U/G. We do not specify how G, the *granule generator*, was obtained. Then the granular information vector is defined as $Inf_{B/G}(u)$ where $u \in U$ is an object representative of a given granule $G_i \in U/G$, and its coordinates correspond to description of the granule associated with u. The value of such coordinate may be quite compound, depending on the granular system we are dealing with. An example of such value could be an interval, a mathematical formula, a fragment of a digital image, or even a subset of objects belonging to a granule with a specific structure assigned.

The granular information vector is introduced in a very general way, as it is meant to be used to describe various types of granules. We want to relate to situation when in the process of creation of granules we end up with a collection of objects (granules, tuples, subsets) that are described by means of information vector consisting of compound, involved attributes that may not be intuitive to use. Creation of such attributes is a consequence of applying a particular granule generator. If we consider, for example, a clustering method for multidimensional domain, then the description would be a set of numbers that describe the clusters, their centers, radii, distributions, and so on.

In the paper we frequently use the notions of quality measure and/or quality criterion. Albeit the measures and criteria that we will refer to are very varied, and not all of them are simply the sets of numerical values, we would like to use a unified way of discussing them. There are two kinds (or rather two levels) of quality measures we are concerned with. In Figure 1 we introduce Q_U as notation for the quality evaluation without any granularity involved. We also introduce $Q_{U/G}$ to denote the corresponding quality evaluation result for system with granularity. The $Q_{U/G}$ represents for us the higher level quality measure, one that judges the granular system on the basis of its performance in the particular application. The other type of measures resides in the part of Figure 1 corresponding to representation. These measures are used to select and control the granule generators. Insofar, most of the investigations concentrated on this (representation) level. In particular, Pedrycz in [13] consider the principle of *justifiable granularity*. In a nutshell, the justifiable granularity principle require granule generators (algorithms for creation of granules) to work in such a way that the resulting granules are strongly supported by experimental evidence (*justified* by original data) and semantically *meaningful* (of sufficient specificity, sound level of detail). This principle mostly pertains the creation of granules, i.e., finding the proper granular representation. It can be materialized in the form of numerical evaluation functions for various types of granular environments, e.g., fuzzy-set-based or interval-based [2].

Our focus presented in this paper extends the above investigations or, in other words, treats their outcomes as one of necessary components of the granular system evaluation framework. One may say that the above-mentioned numerical evaluation functions are heuristic measures that – if appropriately designed and tuned – should anticipate whether a granular

system is going to perform accurately and efficiently in practice. While the approach of Pedrycz ([13]) postulates the creation of granules to be treated as a two-way trade-off between justifiability (support) and meaningfulness (specificity), we want to include more factors, in particular, usefulness and applicability. The optimization of the granular system in our approach is a three-way wager between justifiability, usefulness, and versatility. The versatility is understood by means of the algorithms that will operate on granules in the application layer (see Fig. 1). We would like to have a granular system that nicely balances the effort needed to obtain the result using granules and the (possibly reduced) precision of the final result.

The granules, represented by corresponding granular information vectors $Inf_{B/G}$, are subjected to algorithms for creation of a model or a computational scheme (a computation on granules) that we are looking for (e.g., classification, information retrieval, or query processing). The resulting model is then evaluated with respect to quality, accuracy and applicability. In this presentation, we want to make the case for a systemic approach to quality assessment. In our view the final score should be established by calculating an analog of the empirical risk measure. Such risk-like measure is taken as a summarized expectation for creating a loss (or gain) due to use of particular granular system in a given application. It is quite common, and can be argued as commonsense, to make assessment of the quality of solution by hypothesizing the situations in which the gain/loss can be generated in our system, and then weighting them by the likelihood of their occurrence. The necessary step in this "granular risk assessment" is the calculation of gain (or loss). The proper choice of loss/gain function is a challenging task in itself, but this is the point where we may make a clever use of previously calculated "standard" measures such as accuracy, computational cost, etc.

3 Examples of Granular Systems

To justify that the proposed risk-like measures can be adopted across the wide range of granular systems we provide a set of examples. These examples range from relatively straightforward applications of rough and fuzzy granules in construction of classification systems to very complex construction of data warehouses and building semantic indexes for information search and retrieval engines.

The first and most straightforward example is the granular information system with granules based on simple rough set approach. The granule generation algorithm G is, in the simplest version, the one of selecting a subset of attributes. To have as a result a really useful and versatile granular system we cannot simply take a reduct (see [12]). We have to account for overfitting and relevance. For the purpose of getting better granular system we may apply several that not only select the features (attributes) in a clever way (see [7, 8, 16]), but modify the original information vectors in such a way that it

becomes more natural to represent them as granules. Such approaches can be found in [1, 10, 11] and [15], where instead of attribute values the authors deal with groups, intervals or clusters. With the right selection of granules we may proceed to construction of a quality measure that is an equivalent of empirical risk functional (see [21]), as proposed in [18] and [19].

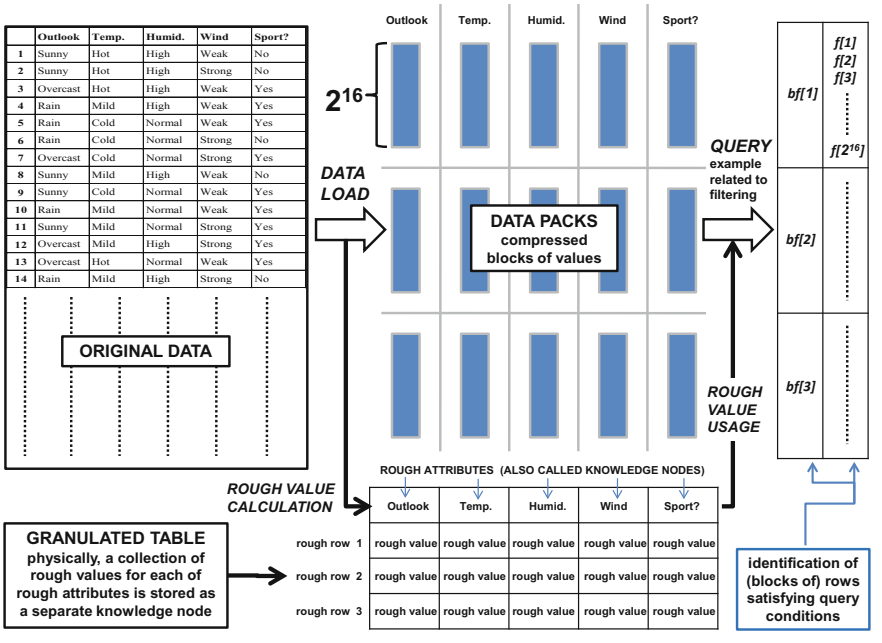


Fig. 2 Infobright's ICE/IEE granular RDBMS engine (see http://www.infobright.org) supports storage and processing of very large data. ICE/IEE automatically creates granulated tables during data load with rows corresponding to the groups of original rows and attributes corresponding to various forms of compact information.

3.1 Granularity in Data Warehousing

In the second example we showcase how the granularity may be incorporated into construction of data warehousing solutions. We rely on Infobright's ICE/IEE ([17],http://www.infobright.org), an RDBMS engine that is meant to support the storage and analytic processing of large amounts of data. One of the defining features of ICE/IEE is that it automatically creates granulated tables during data load. When loading data from an external file to a previously declared table, the rows are decomposed onto their particular attributes' values. After gathering 2^{16} of new rows, ICE/IEE automatically

computes rough values and compresses the blocks of 2^{16} of values corresponding to each of the table's attributes. Rough values and compressed blocks of values are stored on disk. When querying, all applicable rough values are put into memory instead of all corresponding original data values. In case of, e.g., data filtering, rough values are applied in order to quickly exclude the blocks of values that do not satisfy a given SQL condition. The process of loading data (granulating the table) is symbolically presented in Figure 2.

The risk-like quality measures we are looking for in this application should help us in determining the most appropriate ways to structure and process the data, one that will give the best trade-off between quality and speed of answer to the SQL query (see [3],[5],[9]). By using a granulated information during query processing we want to make significant savings on computational effort. If we could retrieve and use the information stored in the database using only the algorithms that work on granular representation, then significant savings are possible, provided that the granule representation is efficient (well optimized, compact, quick).

3.2 Granularity in Semantic Search

The final example comes from the area of semantic information storage, search and retrieval. Within a framework of a larger (SYNAT) project (see [4]) we want to design and implement a solution that will make it possible for a user to search within repositories of scientific information (articles, patents, biographical notes, etc.) using their semantic content. Our prospective architecture for doing that is called SONCA (abbreviation for **S**earch based on **ON**tologies and **C**ompound **A**nalytics).

Ultimately, SONCA should be capable of answering the user query by listing and presenting the resources (documents, Web pages, et caetera) that correspond to it *semantically*. In other words, the system should have some *understanding* of the intention of the query and of the contents of documents stored in the repository as well as the ability to retrieve relevant information with high efficacy. The system should be able to use various knowledge bases related to the investigated areas of science. It should also allow for independent sources of information about the analyzed objects, such as, e.g., information about scientists who may be identified as the stored articles' authors.

The granularity in SONCA model emerges quite naturally when we proceed with construction of the relational data schema aimed at efficient storage and querying of parsed scientific articles, as well as entities corresponding to authors, institutions, references, scientific concepts, et cetera (see upper-right part of Figure 3). An important requirement of the proposed model is to be able to answer a query about all possible entities that may be interesting for users in a well organized and efficient manner. To be able to do that the

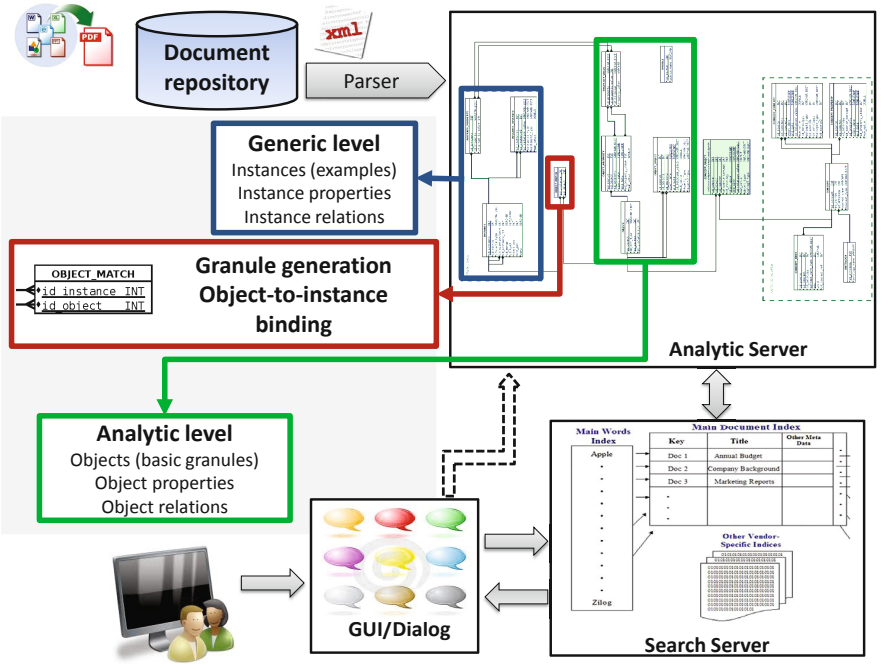


Fig. 3 Granulation in matching [4].

process of *matching*, tantamount to generation of granules of instances, has to be performed.

Matching in the analytical part of SONCA database architecture is the process that takes instances one-by-one and groups them them together if they represent semantically similar objects (entities). A simplest example would be to match all possible appearances of given person's name, regardless of abbreviations, errors, changed case, etc. Another elementary case for matching is the gathering of all occurrences of a given research paper and its citations, regardless of the way the reference was made. In general, the problem of identifying (matching) named entities, especially when they possess an internal structure (as in case of XML documents), is visible in many applications , as discussed in [6].

The result of matching is stored in special table on the database server. This table stores the binding information for instances (records in original data) and granules (dubbed *objects*). The objects may in fact represent much more complicated types of granules than just all occurrences of some person, article, or institution. Ultimately, the fact that two instances are together in one granule should represent their conceptual relationship, i.e., their semantical relatedness.

4 Conclusions

In this paper we have proposed a unifying approach to looking at the quality of a granular systems. We want the quality assessment process for various, possibly very varied, types of granular systems to follow the common pattern. By showcasing several examples of real granular systems we demonstrated, that indeed there are similarities in their composition that call for a systemic approach.

The issues raised in the paper are quite general and to address them to full extent is a challenge for the future. In the particular fields of applications, such as those presented as examples, we hope to be able to improve on some aspects of the underlying granular systems and achieve better, more useful and versatile outcomes.

References

1. Aggarwal, C.C. (ed.): Data Streams - Models and Algorithms. Advances in Database Systems, vol. 31. Springer (2007)
2. Apolloni, B., Pedrycz, W., Bassis, S., Malchiodi, D.: The Puzzle of Granular Computing. SCI, vol. 138. Springer (2008)
3. Beaubouef, T., Petry, F.E.: Uncertainty Modeling for Database Design Using Intuitionistic and Rough Set Theory. Journal of Intelligent and Fuzzy Systems 20(3), 105–117 (2009)
4. Bembenik, R., Skonieczny, Ł., Rybiński, H., Niezgódka, M. (eds.): Intelligent Tools for Building a Scientific Information Platform. SCI, vol. 390. Springer (in print, 2012)
5. Chaudhuri, S., Narasayya, V.R.: Self-tuning database systems: A decade of progress. In: Koch, C., Gehrke, J., Garofalakis, M.N., Srivastava, D., Aberer, K., Deshpande, A., Florescu, D., Chan, C.Y., Ganti, V., Kanne, C.C., Klas, W., Neuhold, E.J. (eds.) VLDB, pp. 3–14. ACM (2007)
6. Dorneles, C.F., Goncalves, R., dos Santos Mello, R.: Approximate data instance matching: a survey. Knowl. Inf. Syst. 27, 1–21 (2011), doi: http://dx.doi.org/10.1007/s10115-010-0285-0
7. Guillet, F., Hamilton, H.J. (eds.): Quality Measures in Data Mining. SCI, vol. 43. Springer (2007)
8. Guyon, I., Elisseeff, A.: An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3, 1157–1182 (2003)
9. Idreos, S., Manegold, S., Kuno, H.A., Graefe, G.: Merging what's cracked, cracking what's merged: Adaptive indexing in main-memory column-stores. PVLDB 4(9), 585–597 (2011)
10. Kreinovich, V.: Towards Faster Estimation of Statistics and ODEs Under Interval, P-Box, and Fuzzy Uncertainty: From Interval Computations to Rough Set-Related Computations. In: Kuznetsov, S.O., Ślęzak, D., Hepting, D.H., Mirkin, B.G. (eds.) RSFDGrC 2011. LNCS (LNAI), vol. 6743, pp. 3–10. Springer, Heidelberg (2011)
11. Nguyen, H.S.: Approximate Boolean Reasoning: Foundations and Applications in Data Mining. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets V. LNCS, vol. 4100, pp. 334–506. Springer, Heidelberg (2006)

12. Pawlak, Z., Skowron, A.: Rudiments of Rough Sets. Information Sciences 177(1), 3–27 (2007)
13. Pedrycz, W.: The Principle of Justifiable Granularity and an Optimization of Information Granularity Allocation as Fundamentals of Granular Computing. Journal of Information Processing Systems 7(3), 397–412 (2011)
14. Pedrycz, W., Skowron, A., Kreinovich, V. (eds.): Handbook of Granular Computing. John Wiley & Sons (2008)
15. Sakai, H., Ishibashi, R., Koba, K., Nakata, M.: Rules and Apriori Algorithm in Non-deterministic Information Systems. In: Peters, J.F., Skowron, A., Rybiński, H. (eds.) Transactions on Rough Sets IX. LNCS, vol. 5390, pp. 328–350. Springer, Heidelberg (2008)
16. Ślęzak, D., Janusz, A.: Ensembles of Bireducts: Towards Robust Classification and Simple Representation. In: Kim, T.-h., Adeli, H., Slezak, D., Sandnes, F.E., Song, X., Chung, K.-i., Arnett, K.P. (eds.) FGIT 2011. LNCS, vol. 7105, pp. 64–77. Springer, Heidelberg (2011)
17. Ślęzak, D., Kowalski, M.: Intelligent Data Granulation on Load: Improving Infobright's Knowledge Grid. In: Lee, Y.-h., Kim, T.-h., Fang, W.-c., Ślęzak, D. (eds.) FGIT 2009. LNCS, vol. 5899, pp. 12–25. Springer, Heidelberg (2009)
18. Szczuka, M.: Risk Assessment in Granular Environments. In: Peters, J.F., Skowron, A., Chan, C.-C., Grzymala-Busse, J.W., Ziarko, W.P. (eds.) Transactions on Rough Sets XIII. LNCS, vol. 6499, pp. 121–134. Springer, Heidelberg (2011)
19. Szczuka, M.S., Skowron, A., Stepaniuk, J.: Function Approximation and Quality Measures in Rough-Granular Systems. Fundamenta Informaticae 109(3), 339–354 (2011)
20. Szczuka, M.S., Ślęzak, D.: Feedforward Neural Networks for Compound Signals. Theoretical Computer Science 412(42), 5960–5973 (2011)
21. Vapnik, V.: Statistical Learning Theory. John Wiley & Sons (1998)
22. Zadeh, L.A.: Toward a Generalized Theory of Uncertainty (GTU) - An Outline. Information Sciences 172(1-2), 1–40 (2005)

Similarity Analysis Based on the Weighted Moving Window for Tissue Characterization of Plaque in Coronary Arteries

Gancho Vachkov and Eiji Uchino

Abstract. This paper is dealing with the problem of tissue characterization of the plaque in the coronary arteries by processing the data from the intravascular ultrasound catheter. The similarity analysis method in the paper is applied in the frame of the moving window approach, which scans all cells in the matrix data from one cross section of the artery. The center-of-gravity model is used for evaluating the dissimilarity between any given pairs of data sets, belonging to pairs of windows. As a computational strategy, the use of weighted values of dissimilarity within the cells belonging to one window is proposed in the paper, rather than simply using an equal mean value for all cells in the window.

The similarity results from each cross section of the artery are displayed as gray scale image, where the darker areas denote the more similar areas to a predefined region of interest. The simulation results from the tissue characterization of a real data set show that the weighted moving window approach gives a sharper resolution of the similarity results that are closer to the real results, compared to the simple mean value approach. This suggests that the weighted moving window approach can be applied to real medical diagnosis.

Keywords: Similarity analysis, Weighted mowing window, Tissue characterization, Intravascular ultrasound, Classification.

1 Introduction

Health condition of the coronary arteries is vital for the normal functioning of the human heart since they supply a fresh blood to the muscular tissue of the heart.

Gancho Vachkov
Graduate School of Science and Engineering, Yamaguchi University, 1677-1 Yoshida, Yamaguchi 753-8512, Japan

Eiji Uchino
Fuzzy Logic Systems Institute (FLSI), 630-41 Kawazu, Iizuka, Fukuoka 820-0067, Japan
e-mail: {vachkov,uchino}@yamaguchi-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 297–307.
springerlink.com

When a gradual build-up of a *plaque* in the inner surface of the artery has occurred, it may cause under some circumstances severe heart diseases (acute coronary syndrome) such as myocardial infarction and angina.

The inner structure of the plaque tissue is directly related to the risk of a heart failure. Here the most important are two types of plaque structure, namely the *lipid* and the *fibrous* structure. A lipid plaque covered by a small and thin fibrous plaque is very likely to break and enter the blood stream, thus creating dangerous blood clots. Therefore it is of utmost importance to analyze the structure of the plaque and find out the so called lipid and fibrous *regions of interest*, abbreviated as Lipid ROI and Fibrous ROI.

The analysis and estimation of the size, shape and location of the Lipid ROI and Fibrous ROI is usually called *tissue characterization* in medical terms, which falls into the research area of pattern recognition and pattern classification.

One of the most frequently used techniques to get reliable information from the coronary artery for further visualization and tissue characterization is the *Intravascular Ultrasound* (IVUS) method [1]. The IVUS method uses a small rotating catheter with a probe inserted into the coronary artery that emits a high frequency ultra-sonic signal to the tissue. The reflected radio-frequency (RF) signal is measured and saved in the computer memory for any further analysis and visualization.

The IVUS is essentially a *tomographic imaging* technology, in which the reflected RF signal is preprocessed to produce a gray-scale image with a circular shape, called a *B-mode image* that is used by medical doctors for observation and analysis of the artery occlusion. One B-mode image corresponds to one *cross-section* of the coronary artery with a given depth-range in all *256* directions (angles) of rotation of the IVUS probe.

The main goal in this paper is not just the data visualization, but rather developing an appropriate method for tissue characterization. Therefore further on we represent the data and the respective results in a rectangular *X-Y* shape, instead of in circular shape. The axis *X* denotes the *angle* (direction) of the IVUS probe within the range of [*0, 255*], while the ordinate *Y* denotes the *depth* of the measurement. i.e. the *distance* between the probe and the current measured signal. The depth-of-interest in our investigations is within the range: [*0, 400*] since any lipid ROI found in the deeper inner areas of the coronary artery is considered as “not so risky”.

A graphical illustration of the matrix-type information obtained by the IVUS probe is presented in Fig. 1.The obtained large size matrix is called RF *matrix* and is further on saved in the computer memory It consists of every single measurement obtained for the IVUS probe for one cross section in the coronary artery.

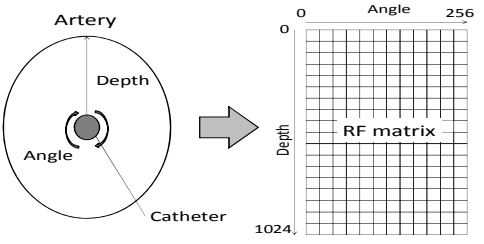


Fig. 1 Information obtained from the IVUS probe for one cross-section of the coronary artery. This is a rectangular X-Y matrix of data with *X* denoting the rotation angle and *Y* denoting the depth of the measurement.

Deep and long term research and data analysis have been done until now [1] – [4] to utilize the information obtained from the IVUS method for a proper tissue characterization of the plaque in the coronary artery. As a result, different classification techniques and algorithms have been proposed, developed and used for various simulations and comparisons. However, currently no "ideal" and easy-to-apply method still exists.

In this paper we propose the use of the moving window approach for *similarity analysis* of the data in the RF matrix, in order to find regions that are most similar to a given (pre-specified) Lipid ROI or Fibrous ROI. The work in this paper is a more detailed and advanced step of our previous work in [4] that also uses the concept of moving window and similarity analysis. In this paper we introduce a moving window with *maximal* overlapping ratio, a new method for similarity analysis and also a new way for calculating the similarity, based on the concept of the weighted moving window.

The rest of the paper is organized as follows. In Section 2 the standard moving window for similarity analysis is explained and in Section 3 the computational details of the new proposed weighted moving window are given. Section 4 explains the details of the center-of-gravity method used for similarity analysis. The experimental results are given in Section 5 and Section 6 concludes the paper.

2 The Standard Moving Window for Similarity Analysis

2.1 The Standard Moving Window Approach

The idea of the standard moving window approach is well known and used in many applications, especially in the research field of image processing and also in other areas that use a large *rectangular* data set as initial information.

First of all, a rectangular matrix of data to be analyzed should be available. In this research we use the RF data matrix, which contains all the reflected signal intensities from one *cross section* of the artery. This matrix consists of all the measured values of the reflected RF signal and has a size of *256* x *Dmax*, where *Dmax = 400* is considered as *sufficient maximal depth* for examination.

The next step is to select the *window size* $N_X \times N_Y$, with "reasonable" values for the *Angle_Range:* $2 \le N_X \le 60$ and for the *Depth_Range:* $5 \le N_Y \le 100$. Obviously these are *problem dependent* parameters that affect the final results from the tissue characterization. In many cases *reasonably* small window size gives better similarity results, while extremely small or extremely large sizes lead to deterioration.

The moving windows approach in fact performs a *scanning* procedure that starts with the first window being located at the *upper-left* corner of the RF matrix. Then this window gradually moves to the right with one-step (*one angle* position) only until the end of the line. After that the scanning is returned to the leftmost position, but shifted with one step (*one depth* position) downwards and continues on this new line. The scanning procedure continues until the last window reaches the bottom-right corner of the RF matrix.

Such scanning procedure ensures that every two neighboring windows have a *maximal overlapping* ratio, since they differ from each other by only one horizontal and one vertical position. This way of movement of the windows is different from our previous moving window approach in [4] where *no overlapping* between the neighboring windows was assumed.

The moving window process with maximal overlapping ratio leads to generating a large number of windows, thus increasing the overall computation time. The total number of the generated windows is: $N_W = 256 \times (D_{\max} - N_Y + 1)$. The total computation time depends not only on the number of the windows, but also on the complexity of the model that is calculated from the data in each window. In order to alleviate the computational burden, we have proposed in Section 4 a simple and easy to calculate representative model, called center-of-gravity model.

2.2 Similarity Analysis by Using the Moving Window

The similarity analysis used in frame of the moving windows approach is basically a method for calculating the difference (dissimilarity) between the structures of two data sets at each scanning step.

The first data set is fixed (constant) and is used as a *reference* data sets (an example) for comparison during the scanning process. These data should be available before the scanning process. They correspond to one typical (and correctly identified) Fibrous or Lipid ROI. It is obvious that the proper identification of such ROI depends on the medical doctor decision and experience.

The other (second) data set used for the similarity is different for each scanning step and is extracted from the respective window W_i, $i = 1, 2, ..., N_W$ for this step.

It becomes clear now that the proposed similarity analysis computation is essentially a *supervised procedure* for decision making, in which the *Dissimilarity Degree* DS at each step is calculated between one reference data set and one *unknown* set, belonging to the current window.

The value of DS is usually bounded: $DS_i \in [0,T]$ and shows *how close* is the data structure from the current window W_i, $i=1,2,...,N_W$ to the data structure in the predefined ROI. A value of dissimilarity, close to zero suggests that the two data sets are very similar while a bigger (closer to *T*) value stands for a bigger difference (bigger discrepancy) between the two data sets.

For the purpose of *fair* quantitative comparison between the data sets, the dissimilarity value is often *normalized* as: $DS_i \in [0,1]$.

It is clear that the calculated value of dissimilarity depends on the assumed model for describing the structure of the data set in each window. A simple and yet effective model for similarity analysis is presented in Section 4.

During the moving window process, the similarity value DS_i is calculated many times, namely for each current position of the window W_i. Then the main question is *where* (at which location) to assign the currently calculated value of DS_i? This problem arises because all $N_X \times N_Y$ data in the current window W_i have been used for calculation of the dissimilarity.

The simplest and reasonable decision is to assign the *same dissimilarity* degree DS_i to *all* coordinates (all *cells*) within the current window W_i that have been used in the calculations. In order to keep in a memory all these values, we create a *new* rectangular matrix called *Dissimilarity Matrix* DM, with the same dimension as the Data Matrix RF, i.e. *245* x *Dmax*.

It is easy to realize that each data item (each cell) in the original data matrix RF will be visited not only ones, but many times by the moving window. If *N* denotes the number of all visits of a given cell at location $\{i,j\}$; $i \in [0,255]$; $j \in [0,Dmax]$ by a moving window with size $N_X \times N_Y$, then this number is calculated as:

$$\begin{aligned} N &= (j+1)N_Y, \ \ if \ \ 0 \le j \le N_Y - 2; \\ N &= N_X \times N_Y, \quad if \ \ N_Y - 1 \le j \le D_{\max} - N_Y + 1; \\ N &= (D_{\max} - j + 1)N_Y, \ \ if \ \ D_{\max} - N_Y + 2 \le j \le D_{\max}; \end{aligned} \tag{1}$$

The DM matrix is actually an *additive matrix* that accumulates all the calculated values for the dissimilarity degrees at the same location $\{i,j\}$, as follows:

$$dm_{ij} \leftarrow dm_{ij} + DS_k, \ \ k=1,2,...,N \tag{2}$$

The *final dissimilarity* degree for each $\{i,j\}$ location can be taken in different ways, but in the in the simplest case of the standard moving window, we take it as a *mean value* of the accumulated similarity degrees in DM, namely:

$$dm_{ij} = dm_{ij} / N, \ \ i \in [0,255]; \ j \in [0, D_{\max}] \tag{3}$$

All final dissimilarity degrees (3) are saved in the DM matrix. Further on, they can be displayed in the following two ways for a final decision making:

- *Hard Decision* making with a user-defined *threshold Th* for separation of all the values into *2* crisp classes: *Similar* (with $dm_{ij} \le Th$) and *Not Similar* (with $dm_{ij} > Th$) to the pre-specified region of interest, such as *Lipid* ROI or *Fibrous* ROI. Then the *Similar* only class is displayed as a *crisp* dark image to the medical doctor for his final decision. Obviously the threshold selection is not a trivial task that could lead sometimes to ambiguous results.
- *Soft Decision* making. It is a kind of *fuzzy* way of displaying the results, in which all calculated values in (3) are visualized as a *gray scale* image with varying intensities. The similarity values closer to *zero*, corresponding to *very similar* areas, are displayed as *darker areas.* The values closer to *one*, corresponding to *less similar* areas are displayed as *brighter areas.* This gives additional information to the medical doctor in taking his final decision.

3 The Weighted Moving Window for Similarity Analysis

The idea of taking the dissimilarity degree in (3) as a *mean value* of all calculated dissimilarities in one window has a certain drawback. It is that the calculated *true* value in the center depends *equally* on all neighboring dissimilarities, even on those that are far from the center. As a result such assumption may lead to a serious deviation from the actual dissimilarity.

In order to make more precise calculation of the dissimilarities at each cell of the window, we propose here the idea of the *weighted moving window*. The essential point is that the dissimilarity in each cell $[i, j]$of the matrix DM is now *weighted* between *0* and *1* according to its *distance* to the center $[i_0, j_0]$of the window, as shown in the next equation.

$$w_{ij} = exp^{-\frac{|i-i_0|+|j-j_0|}{2\sigma^2}} \in [0,1];\ i=1,N_X;\ j=1,N_Y \tag{4}$$

As seen from (4), a Gaussian function with *Manhattan* (City Block) *distance* and a predetermined *spread* σ is assumed here to evaluate the amount of the weight for each cell in the window. For example, if the window size is 5 x 7, its center location will be [3,4].Then, the Manhattan distance between the center and a cell located at [4,2].will be 3.

The calculated weight by (4) is used to calculate the weighted dissimilarity that will be added to each cell of the matrix DM, in a similar way as in (2), namely:

$$dm_{ij} \leftarrow dm_{ij} + w_{ij} DS_k,\ k=1,2,...,N \tag{5}$$

The *final dissimilarity* degree for each {*i,j*} location will be calculated now as *weighted average* of the accumulated dissimilarities, taking into account the sum of all weights, namely:

$$dm_{ij} = dm_{ij} \Big/ \sum_{n=1}^{N_X} \sum_{k=1}^{N_Y} w_{nk}\,,\ i=1,N_X;\ j=1,N_Y \tag{6}$$

Thus all the final dissimilarities in (6) will be normalized between *0* and *1*.

An illustrative example of the Gaussian function from (4) is given in Fig.2.

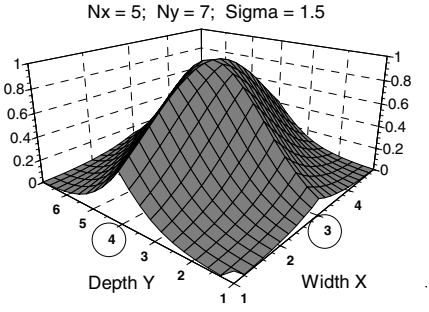


Fig. 2 Example of the Gaussian weight function from (4) used for calculating the weights of the dissimilarities in (5) in the case of a chosen window with size *5* x *7*.

4 The Center-of-Gravity Model Used for Similarity Analysis

In order to calculate the similarity degree between any two data sets, we need a *model* that describes appropriately the structure of the data set. Then, as an initial step, we calculate (once only) the two *Reference Models* RM_L and RM_F for the predefined *Lipid* ROI and *Fibrous* ROI., by using the respective available data: M_L and M_F. After that a *repetitive* model calculation is performed for each of the moving windows that contain the same number of $M_W = N_X \times N_Y$ data.

The Center-of-Gravity (COG) model, proposed and used in this paper produces a simple, but still representative estimation of the structure of the extracted data from each window. The COG model uses just two parameters with clear physical meaning, namely the *Center-of-Gravity* (CG) and the *Standard Deviation* (SD). The presumption is that data sets with different structures have different values of CG and SD, so the amount of the difference between the two parameters can be used to evaluate the similarity degree between them.

As initial information for constructing the RF matrix, we use the raw (actual) reflected RF signal with a simple *preprocessing*, as shown in Fig. 3, namely taking the absolute vale of the RF signal after subtracting the central (stationary) value of *2000* from it.

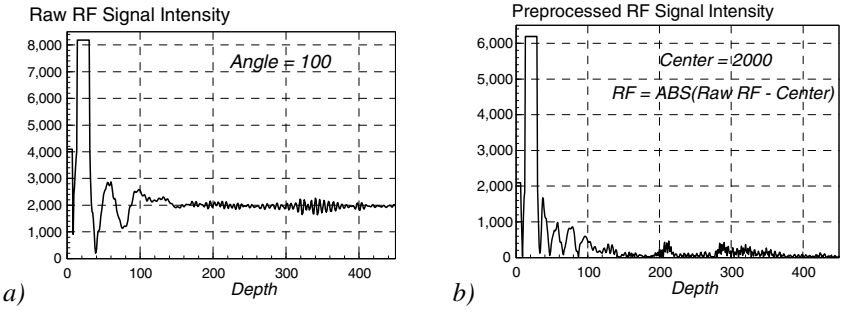


Fig. 3 The Intensity of the Raw (a) and the Preprocessed (b) reflected RF signal at one fixed angle (*100*) of the IVUS catheter and all depths, starting from *0* to *450*.

Let *M* denotes the number of data, i.e. the number of the preprocessed RF signal intensities R_i, $i=1,2,...,M$ from a given ROI or a given window W. Then the two model parameters CG and SD are calculated as follows:

- The *Center-of-Gravity* of the model is simply calculated as the *mean* value of the one-dimensional RF signal:

$$CG = \sum_{i=1}^{M} R_i / M \tag{7}$$

-The *Standard Deviation* is calculated as:

$$SD = \sqrt{\sum_{i=1}^{M} (R_i - CG)^2 \Big/ (M-1)} \tag{8}$$

For example, the calculated values of CG and SD for the Fibrous ROI and the Lipid ROI, extracted from the one experimental RF data matrix are as follows:

Fibrous ROI: CG = *47.565*; SD = *34.344*;
Lipid ROI : CG = *45.375*; SD = *17.045*.

For calculating the *normalized dissimilarity* DS between two COG models, namely the reference model *Mo* and the current window model *Mi*, we propose in this paper the following formula:

$$DS_i = DS(M_0, M_i) = 1 - exp^{-\frac{(CG_0 - CG_i)^2}{2\sigma_C^2}} exp^{-\frac{(SD_0 - SD_i)^2}{2\sigma_S^2}} \in [0,1] \tag{9}$$

Here σ_C and σ_S denote the predefined *width* for the *CG* and *width* for the *SD*, respectively. It is important to note that we are able to control the process of dissimilarity analysis by changing in appropriate way the tuning parameters σ_C and σ_S. For example, a smaller selected value of σ_S , i.e. a narrower width σ_S will give bigger importance to the standard deviation DS of the extracted data, than to

the center-of-gravity CG. Therefore the proper choice of these two tuning parameters (widths) is left to the user.

5 Simulation Results from Tissue Characterization

5.1 Simulation Details and Conditions

The above described moving window approach for similarity analysis was used for tissue characterization of several sets of real data, in the form of respective RF matrices, each of them with *X-Y* size: *256* x *400* (the maximal depth: *Dmax = 400*). For each matrix, the respective *Lipid* ROI and *Fibrous* ROI have been properly identified and marked by a medical doctor through a microscopic analysis. These ROI data were used for creating the *Reference* Lipid and Fibrous *models* for similarity analysis and also for testing and analyzing the correctness of the simulation results.

In the simulations, a moving window with size *21* x *31* was chosen, which generated *94976* windows in total, used for similarity analysis. Despite this large number, the calculations were fast, due to the simple structure of the proposed COG model. The CPU time for one RF data set was about 5 sec on a computer with 3.0 GHz *Intel 4* CPU unit. This suggests that the proposed tissue characterization method, if "accurate enough", could be applied in almost *real-time* mode.

5.2 Tissue Characterization Results from the Similarity Analysis

In the simulations we used one RF data set, corresponding to one cross section of the artery, for which the size, location and boundaries of the *Fibrous* ROI and the *Lipid* ROI were properly diagnosed by the doctor. The data extracted from these two ROIs were used to calculate the respective *Reference* COG models. Then the two reference models were used for similarity analysis by using the *Soft* decision between *0* and *1*, with pre-selected widths in (9), as follows: $\sigma_C = 8.0$ and $\sigma_S = 5.0$. Such settings put a bigger priority to the standard deviation (the *roughness* of the data) than to the CG (the *mean* value of the data). The user-defined value for the spread in (4) was: $\sigma = 2.5$.

The results from the similarity analysis, based on the proposed weighted moving window are shown in Fig. 4 for the case of Fibrous ROI and in Fig. 5 for the case of Lipid ROI. The results, based on the weighted moving window, with the dissimilarity calculated by (6), are depicted in Fig. 4*a*,4*b* and Fig. 5*a*,5*b,* while Fig. 4*c* and Fig. 5*c* show the results from the standard moving window, where the dissimilarity is taken simply as the mean value (3). It is easy to notice that the weighted moving window approach in Fig. 4*a* and Fig. 5*a* produces *sharper images* with higher resolution, which makes the final human decision easier.

Fig.4*b* and Fig. 5*b* are augmentation of Fig. 4*a* and Fig. 5*a*, where the locations of the pre-specified Fibrous ROI and Lipid ROI are added, in order to validate the correctness of the tissue characterization by our proposed method.

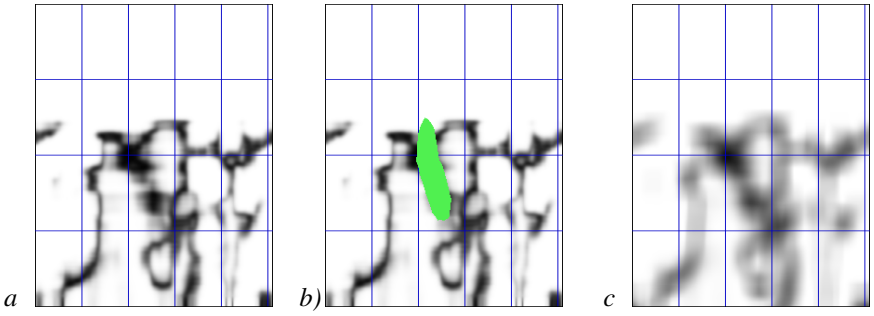


Fig. 4 Tissue characterization results for the case of Fibrous ROI; (a) and (b) are the weighted moving window results; (c) is the standard moving window result.

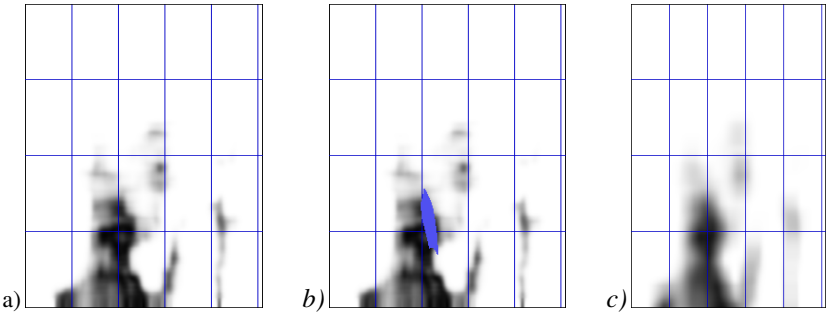


Fig. 5 Tissue characterization results for the case of Lipid ROI; (a) and (b) are the weighted moving window results; (c) is the standard moving window result.

The characterization results show that the proposed similarity analysis method has identified much larger areas as "belonging to" the *Fibrous* ROI and *Lipid* ROI, than the actually two ROIs, identified by the medical doctor. There are several different reasons for such discrepancy between the human and computer results.

One of them is that in reality there could be *several* (multiple) *Fibrous* and *Lipid* ROIs within the examined cross section, however the doctor has diagnosed and marked only one (a *typical*) example of a Fibrous ROI and a Lipid ROI. Another reason could be the *heuristic* (not optimal) selection of the window size $N_X \times N_Y$, as well as the widths σ_C and σ_S. all these are *problem dependent* parameters that have a strong influence on the final tissue characterization results.

6 Concluding Remarks

We proposed in this paper a general Moving Window computational approach with maximal overlapping ratio for tissue characterization of coronary arteries. This approach uses data obtained from the IVUS catheter and allows implementation of different methods and models for similarity analysis. One of them, the

Center-of-Gravity based model was proposed and used in the paper. The moving window approach has been applied in two versions, namely the standard and the weighted mowing window. They differ by the way of calculating the dissimilarity degree DS for the cells in each window. In the standard moving window the average is taken and applied to all cells. In the weighted moving window, a Gaussian function with Manhattan distance is used to assign weights between *0* and *1* of the calculated DS for all cells in the window.

A convenient and practical *soft decision* making for visualization of the similarity results is used in the paper. In this kind of decision all dissimilarities create a gray-scale image, where the *darker* areas represent the *more similar* areas to the predetermined region of interest ROI.

The simulation results by using a real set of data from the IVUS catheter show positive, but still "not perfect" results. They detect successfully large part of the actual Lipid and Fibrous ROI, but also show some other areas, as "very similar" to those ROI.

Possible further improvements of the proposed weighted moving window approach include optimization of the window size and the widths σ_C and σ_C used for decision making. Other ways for improvement of the tissue characterization accuracy are considered now by applying other, more precise models for similarity analysis and other calculation methods for the dissimilarity degree.

Acknowledgment. This work was supported by the Grant-in-Aid for Scientific Research (B) of the Japan Society for Promotion of Science (JSPS) under the Contract No. 23300086.

References

1. De Korte, C., Hansen, H.G., Van der Steen, A.F.: Vascular ultrasound for athero sclerosis imaging. In: Interface FOCUS, vol. 1, pp. 565–575 (2011)
2. Uchino, E., Suetake, N., Koga, T., Kubota, R., et al.: Fuzzy Inference-based plaque boundary extraction using image separability for IVUS image of coronary artery. Electronic Letters 45(9), 451–453 (2009)
3. Koga, T., Furukawa, S., Uchino, E., Suetake, N.: High-Speed Calculation for Tissue Characterization of Coronary Plaque by Employing Parallel Computing Techniques. Int. Journal of Circuits, Systems and Signal Processing 5(4), 435–442 (2011)
4. Vachkov, G., Uchino, E.: Similarity-based method for tissue characterization of coronary artery plaque. In: CD-ROM, Proc. 24th Annual Conference on Biomedical Fuzzy Systems Association (BMFSA), October 20-29, pp. 257–260 (2011)

Simultaneous Optimization of Customer Satisfaction and Cost Function by Nature Inspired Computing

János Botzheim and Péter Földesi

Abstract. When optimizing multi-dimensional non-linear problems the optimal solution in technical terms can be found by heuristic methods. It seems that human thinking does not work properly with the mathematical processing: human decision-makers tend to reject options that represent extreme values in the set of parameters and are not able to handle many system parameters at the same time. The paper investigates the best possible way for modeling the human thinking, comparing bacterial memetic algorithm and particle swarm optimization in fuzzy environment.

1 Introduction

When optimizing multi-dimensional non-linear problems very often heuristic algorithms are applied. The comparison of different approaches and methods are presented in the literature and many reference works highlight the advantages and disadvantages of the algorithms, including the questions of efficiency, convergence, running time and so on [1, 4, 5]. The optimal solution in technical terms can be found with these methods in most of the cases, but it seems that human thinking does not work properly with the mathematical processing: human decision-makers tend to reject options that represent extreme values in the set of parameters. On the other hand human decision-makers are not able to handle many system parameters at the same time, mainly when two ore more objectives and/or restriction functions are applied then either sub-optimal solutions or multi-objective approaches are used. The question is what the best way is for modeling the human thinking,

János Botzheim
Department of Automation, Széchenyi István University, 1 Egyetem tér, Győr, 9026, Hungary
e-mail: botzheim@sze.hu

Péter Földesi
Department of Logistics and Forwarding, Széchenyi István University, 1 Egyetem tér, Győr, 9026, Hungary
e-mail: foldesi@sze.hu

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 309–318.
springerlink.com

so in this paper the Kano non-linear and subjective quality model is investigated in order to compare the possible optional techniques. The bacterial memetic algorithm and the particle swarm optimization are compared in fuzzy environment, and in the numerical example the results show the real difference between the methods, and the trade-off can be clearly seen.

2 Kano's Quality Model

Questions of budgeting is not a key element of Kano's original model [10], but we can reasonably assume, that improving the level of technical attributes requires extra costs, so for each technical attribute a cost function can be set. The general target is to achieve the maximum economic result with the minimum use of resources, that is to maximize customer satisfaction with the minimum cost. The task can be mathematically formulated in two ways:
(A1) maximize overall satisfaction (S) not exceeding given cost limit (C)
or
(A2) achieve given overall satisfaction (S) with minimum cost (C).
(B) Increase customer satisfaction and decrease the cost at the same time.

Let

$$S_i(x_i) = b_i + a_i \cdot x_i^{\widetilde{\beta}_i} \quad i = 1,2,\ldots,n \tag{1}$$

be the customer satisfaction generated by technical attribute x_i, where

- $0 \leq x_i \ll \infty$ is a real variable
- $a_i > 0$ is a real constant
- $\widetilde{\beta}_i$ is a triangular shaped fuzzy number with (β_{iL}, β_{iR}) support and β_{iC} core value, thus $\widetilde{\beta}_i = (\beta_{iL}, \beta_{iC}, \beta_{iR})$.
- b_i is a constant
- n is the number of technical attributes considered in the designing process.

Furthermore let

$$C_i(x_i) = f_i + \widetilde{v}_i \cdot x_i \quad i = 1,2,\ldots,n \tag{2}$$

be the cost of manufacturing technical attribute at level x_i, $f_i \geq 0$ is a real constant and $\widetilde{v}_i$ is a triangular shaped fuzzy number with (v_{iL}, v_{iR}) support and v_{iC} core value, thus $\widetilde{v}_i = (v_{iL}, v_{iC}, v_{iR})$. Let

$$S = \sum_{i=1}^{n} b_i + a_i \cdot x_i^{\widetilde{\beta}_i} \tag{3}$$

be the overall satisfaction and

$$C = \sum_{i=1}^{n} f_i + \widetilde{v}_i \cdot x_i \tag{4}$$

be the total cost, according to the "fixed costs – variable costs" methodology. The reason for using fuzzy numbers in the exponent part of S_i and in describing the cost is that these values are uncertain in most of the cases [7, 8]. The fuzzy exponent

calculation in S_i can be done by a simple parametric representation as presented in [2]. The general formula for (A) is:

- (A1) Let $\sum_{i=1}^{n} S_i(x_i) \rightarrow max$, subject to $\sum_{i=1}^{n} C_i(x_i) \leq C_0$, where C_0 is a given constant.
- (A2) Let $\sum_{i=1}^{n} C_i(x_i) \rightarrow min$, subject to $\sum_{i=1}^{n} S_i(x_i) \geq S_0$, where S_0 is a given constant.

Case (A1) was investigated in [8], and case (A2) was analyzed in [7]. In this paper case (B) is investigated in three different forms:

- (B1) Let $\frac{\sum_{i=1}^{n} S_i(x_i)}{\sum_{i=1}^{n} C_i(x_i)} \rightarrow max$.
- (B2) Let $\sum_{i=1}^{n} \frac{S_i(x_i)}{C_i(x_i)} \rightarrow max$.
- (B3) Let $\sum_{i=1}^{n} S_i(x_i) \rightarrow max$ and $\sum_{i=1}^{n} C_i(x_i) \rightarrow min$ simultaneously.

In all cases $0 \leq x_i \leq x_{iU}$, where x_{iU} is the upper bound of the search space in the i-th dimension.

3 Nature Inspired Computing

Nature inspired algorithms are often suitable for global optimization of even nonlinear, high-dimensional, multi-modal, multi-objective, and discontinuous problems. They are mimicking phenomena that can be found in the nature. In this paper two nature inspired computing techniques are applied, the bacterial memetic algorithm and the particle swarm optimization method. In our case one individual (bacterium or particle) is one vector in the search space, thus the individual contains the coordinates of the position described by real numbers. The length of the individual is constant and equals to the number of dimensions of the search space.

3.1 Bacterial Memetic Algorithm

Bacterial Evolutionary Algorithm (BEA) [17] uses two operators, the bacterial mutation and the gene transfer operation. The bacterial mutation operation optimizes the chromosome of one bacterium, the gene transfer operation allows the transfer of information between the bacteria in the population. Local search approaches might be useful in improving the performance of the basic evolutionary algorithm, which may find the global optimum with sufficient precision in this combined way. Combinations of evolutionary and local-search methods are usually referred to as memetic algorithms [15]. A new kind of memetic algorithm based on the bacterial approach is the bacterial memetic algorithm (BMA) [3]. The algorithm consists of four steps. First, a random initial population with N_{ind} individuals has to be created. Then, bacterial mutation, local search, and gene transfer are applied, until a stopping criterion (number of generations, N_{gen}) is fulfilled. In case of continuous problems gradient-based method can be applied as local search. The Levenberg-Marquardt method was proposed as local search technique in the original version of BMA [3] and it is ap-

plied in this paper, too. In the initial population N_{ind} random individual (bacterium) are created.

3.1.1 Bacterial Mutation

Bacterial mutation is applied to all bacteria one by one. First, N_{clones} copies (clones) of the bacterium are created. Then, a random segment of length l_{bm} is mutated in each clone except one clone which is left unmutated. After mutating the same segment in the clones, each clone is evaluated. The clone with the best evaluation result transfers the mutated segment to the other clones. These three steps operations (mutation of the clones, selection of the best clone, transfer of the mutated segment) are repeated until each segment of the bacterium has been mutated once. At the end, the best clone is kept as the new bacterium and the other clones are discharged.

3.1.2 Levenberg-Marquardt Method

The Levenberg-Marquardt method is a gradient based optimization technique proposed originally by Levenberg [13] and Marquardt [14] for least-square estimation of non-linear parameters. Our task is to find the minimum value of a function $f(\mathbf{x})$. The method is applied to all bacteria one by one. For a given bacterium $\mathbf{x}[k]$ in the k-th iteration step the update vector is:

$$\mathbf{s}[k] = -[J(\mathbf{x}[k]) \cdot J(\mathbf{x}[k])^T + \gamma[k] \cdot \mathbf{I}]^{-1} \cdot J(\mathbf{x}[k]), \tag{5}$$

where $J(\mathbf{x}[k])$ is the gradient vector of $f(\mathbf{x})$ at $\mathbf{x}[k]$, γ is a parameter initially arbitrary set to any positive value ($\gamma[1] > 0$) and $\mathbf{I}$ is the identity matrix. We approximate the derivatives in $J(\mathbf{x}[k])$ by finite differences. After the update vector was computed we calculate the so-called trust region, $r[k]$ as follows: $r[k] = \frac{f(\mathbf{x}[k]+\mathbf{s}[k]) - f(\mathbf{x}[k])}{J(\mathbf{x}[k])^T \cdot \mathbf{s}[k]}$. The value of parameter γ is adjusted dynamically depending on the value of $r[k]$:

- If $r[k] < 0.25$ then $\gamma[k+1] = 4\gamma[k]$
- If $r[k] > 0.75$ then $\gamma[k+1] = \gamma[k]/2$
- Else $\gamma[k+1] = \gamma[k]$

If $f(\mathbf{x}[k] + \mathbf{s}[k]) < f(\mathbf{x}[k])$ then $\mathbf{x}[k+1] = \mathbf{x}[k] + \mathbf{s}[k]$, else $\mathbf{x}[k+1] = \mathbf{x}[k]$. If the stopping condition ($||J(\mathbf{x}[k])|| \leq \tau$) is fulfilled or a predefined maximum iteration number is reached then the algorithm stops, otherwise it continues with the $(k+1)$-th iteration step. The search direction varies between the Newton direction and the steepest direction, according to the value of γ. If $\gamma \to 0$, then the algorithm converges to the Newton method, if $\gamma \to \infty$, then it gives the steepest descent approach.

3.1.3 Gene Transfer

First, the population must be sorted and divided into two halves according to their evaluation results. The bacteria with better evaluation are called superior half, the bacteria with worse evaluation are referred to as inferior half. Then, a source

bacterium is randomly chosen from the superior half and a destination bacterium from the inferior half. A segment of length l_{gt} from the source bacterium is randomly chosen and this segment overwrites the same segment of the destination bacterium. The above steps (sorting the population, selection of the source and destination bacteria, transfer the segment) are repeated N_{inf} times, where N_{inf} is the number of "infections" per generation.

3.2 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population based stochastic optimization technique inspired by social behavior of bird flocking or fish schooling [11, 12]. In these methods the individuals try to find better and better places by exploring their environment led by their own experiences and the experiences of the whole community. Each particle is associated with a position in the search space which represents a solution for the optimization problem. Each particle remembers the coordinates of the best solution it has achieved so far. The best solution achieved so far by the whole swarm is also remembered. The particles are moving towards in the direction based on their personal best solution and the global best solution of the swarm. The positions of the particles are randomly initialized. The next position of particle i is calculated from its previous position $\mathbf{x}_i(t)$ by adding a velocity to it:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1). \tag{6}$$

The velocity is calculated as:

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + c_1 r_1 [\mathbf{p}_i(t) - \mathbf{x}_i(t)] + c_2 r_2 [\mathbf{g}(t) - \mathbf{x}_i(t)], \tag{7}$$

where w is the inertia weight, c_1 and c_2 are parameters, r_1 and r_2 are random numbers between 0 and 1. The second term in equation (7) is the cognitive component containing the best position remembered by particle i ($\mathbf{p}_i(t)$), the third term in the equation is the social component containing the best position remembered by the swarm ($\mathbf{g}(t)$). The inertia weight represents the importance of the previous velocity, while c_1 and c_2 parameters represent the importance of the cognitive and social components, respectively. The algorithm becomes stochastic because of r_1 and r_2. After updating the positions and velocities of the particles the vectors describing the particles' best positions and the global best position have to be updated, too. If the predefined iteration number ($N_{iteration}$) is reached then the algorithm stops. The number of particles ($N_{particles}$) is also a parameter of the algorithm.

3.3 Multi-Objective Optimization

Many real world problems have not only one objective to be optimized but two or more sometime conflicting objectives. These multi-objective problems have usually no single best solution but a so-called Pareto set, which contains equally good

solutions. A solution is Pareto optimal if there exists no such a solution which could decrease one objective without increasing at least one other objective. Multi-objective evolutionary algorithms have a huge literature with lots of applications [4, 5]. Several different multi-objective PSO methods have also been proposed [6, 16]. In this paper the multi-objective bacterial memetic algorithm (MOBMA) is proposed and a novel multi-objective PSO (MOPSO) based on the same inspiration. We apply these methods for solving the multi-objective (B3) problem presented in Section 2.

3.3.1 Multi-Objective Bacterial Memetic Algorithm

In bacterial mutation the problem is how to select the best clone after the mutation of the clones. The selected clone is the non-dominated one. If there are more than one non-dominated clone then we select that clone which dominates the most number of other clones. If there are still more than one clone according to this measure then we select from this set randomly.

The application of Levenberg-Marquardt algorithm for multi-objective problems is proposed in [18]. The idea is to use the gradient-based algorithm for each objective one by one as the problem were a single objective problem. The same concept was suggested for any local search techniques in multi-objective problems in [9].

In the gene transfer the problem is how to order the bacteria in the population in order to spread the information of the good individuals. We use the same technique as in the bacterial mutation. We rank the individuals according to the dominance. First, we rank the non-dominated bacteria according to how many bacteria they dominate. After the non-dominated individuals the dominated ones are ranked also according to the number of bacteria they dominate.

By the above operations the diversity of the population is guaranteed and there will be more non-dominated individuals in the population simultaneously.

3.3.2 Multi-Objective Particle Swarm Optimization

Although several MOPSO algorithm have been proposed [5, 6, 16] we present a novel one based on the same concept as MOBMA. For the updating of the local best position we apply the technique proposed in [6, 16]. We update the local best position of a particle if the new position of the particle dominates its best position or if they are non-comparable to each other.

The finding of the leader particle – the global best particle – is the same problem as finding the best clone after mutation in MOBMA. Among the non-dominated particles the leader will be that particle which dominates the most number of other particles. The previous global best particle also takes part in the dominance calculation. If there are more particles with the same dominance measure then we decide randomly as in case of bacterial mutation in MOBMA.

4 Numerical Example

Three different problems (B1), (B2), and (B3) presented in Section 2 are investigated with BMA and PSO, on three different α levels $\alpha = 0$, $\alpha = 0.5$, and $\alpha = 1$. The parameters of the function used in the simulations are presented in Table 1 [7]. The number of dimensions is 15, and $\forall i\ x_{iU} = 100$, thus $\forall i\ 0 \leq x_i \leq 100$. Table 2 shows the parameters of BMA, and Table 3 illustrates the parameters of PSO. After some trials, these parameter settings seemed optimal. For fair comparison, by applying these parameter settings, the running time of the algorithms is similar in each case. The results are presented in Table 4 and the corresponding solutions are shown in Table 5. In Table 4 the "dominated" column means the number of dominated bacteria in case of MOBMA, while in case of MOPSO the global best particle is displayed.

Table 1 Function Parameters

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
b	50	40	30	20	10	20	20	20	20	20	-10	-8	-8	-6	-5
a	4.6	4.4	4.2	4.9	6	10	10	8	7	10	20	10	22	15	20
β_L	1.15	1.25	1.3	1.2	1.3	0.9	0.9	0.9	0.9	0.9	0.7	0.65	0.6	0.5	0.4
β_C	1.25	1.3	1.35	1.4	1.45	1	1	1	1	1	0.75	0.7	0.65	0.6	0.55
β_R	1.3	1.4	1.4	1.45	1.5	1.1	1.1	1.2	1.2	1.1	0.85	0.75	0.75	0.65	0.6
f	20	30	40	50	50	10	20	10	20	10	10	10	5	5	5
v_L	9	10	13	15	17	8	8	6	7	8	6	6	4	4	3
v_C	10	12	14	16	18	9	9	9	9	9	8	7	6	5	4
v_R	13	13	15	19	21	10	10	10	10	10	9	8	7	7	6

Table 2 BMA Parameters

	N_{gen}	N_{ind}	N_{clones}	l_{bm}	N_{inf}	l_{gt}	γ init.	max LM iter.	τ
(B1)	100	50	12	1	12	1	1	8	0.0001
(B2)	80	50	12	1	12	1	1	8	0.0001
(B3)	50	50	12	3	12	1	1	8	0.0001

Table 3 PSO Parameters

	$N_{iterations}$	$N_{particles}$	w	c_1	c_2
(B1)	1000	2000	1	0.5	0.5
(B2)	500	2000	1	0.5	0.5
(B3)	100	200	1	0.5	0.5

Table 4 Results

	problem	algorithm	α	S	C	S/C	$\sum_{i=1}^{n} S_i/C_i$	dominated
(1)	(B1)	BMA	1	5050.52	2181.22	2.315455	–	–
(2)	(B1)	BMA	0.5	4853.09	2183.5	2.222614	–	–
(3)	(B1)	BMA	0	4680.87	2186.35	2.140952	–	–
(4)	(B1)	PSO	1	8101.31	3837.77	2.110941	–	–
(5)	(B1)	PSO	0.5	7696.1	3835.48	2.006554	–	–
(6)	(B1)	PSO	0	6766.01	3568.74	1.895909	–	–
(7)	(B2)	BMA	1	13165.5	7451.91	1.766728	24.170449	–
(8)	(B2)	BMA	0.5	12861.93	7454.27	1.725444	24.034948	–
(9)	(B2)	BMA	0	13740.31	8325.44	1.650400	24.204526	–
(10)	(B2)	PSO	1	13165.84	7452.15	1.766717	24.170450	–
(11)	(B2)	PSO	0.5	13628.13	8320.72	1.637855	23.921832	–
(12)	(B2)	PSO	0	16129.43	10700.37	1.507372	22.556622	–
(13)	(B3)	BMA	1	11338.86	7529.61	1.505903	–	15
(14)	(B3)	BMA	1	10602.36	7303.91	1.451601	–	15
(15)	(B3)	BMA	1	10625.95	7349.64	1.445778	–	15
(16)	(B3)	BMA	0.5	11565.06	7445.37	1.553322	–	15
(17)	(B3)	BMA	0	10459.9	6754.44	1.548596	–	15
(18)	(B3)	PSO	1	10236.08	5762.98	1.776178	–	global best
(19)	(B3)	PSO	0.5	10802.42	6128.33	1.762702	–	global best
(20)	(B3)	PSO	0	10413.97	5828.33	1.786785	–	global best

5 Conclusions

Considering the results presented in Tables 4 and 5 the following general tendencies can be detected:

The best solutions subject to the formal objective function (that is subject to the target to maximize overall satisfaction with the minimum cost) can be achieved by applying single objective BMA that can find the best n-dimensional $\mathbf{x}$. Note, that in the optimal $\mathbf{x}$ some of the components are (practically) zero, the others are at the upper limit. (B1 problem).

When considering real human thinking (the behavior of possible customers of a given product) we can assume, that the decision makers are not able to calculate the objective function of an $n = 15$ dimensional problem, and in order to simplify the problem either they try to find the sub-optimal points or they follow a multi-objective search process resulting in a Pareto set. (B2 and B3 problems).

The formal objective function values of sub-optimal search and multi-objective search are smaller than the original solution (B1) which is the general expectation with no doubt. The real interesting finding of this paper are that the sub-optimal search can give better results than the multi-objective heuristics in formal terms (higher objective funcion values: e.g. row (7) and row (10) in Table 4.) but from practical point of view only the multi-objective BMA results are acceptable (rows

Table 5 Solutions

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
(1)	0	0	0	0	100	0	0	0	0	0	0.51	0.06	1.29	0.45	1.27
(2)	0	0	0	0	100	0	0	0	0	0	0.62	0.07	1.48	0.5	1.28
(3)	0	0	0	0	100	0	0	0	0	0	0.74	0.08	1.8	0.54	1.25
(4)	0	0	0	100	100	0	0	0	0	0	0	0	1.67	0	0
(5)	0	0	0	100	100	0	0	0	0	0	0.93	0	0	0	0
(6)	0	0	100	0	100	0	0	0	0	0	0	0	0	0	1.63
(7)	0	100	100	100	100	0	99.96	0	0	0	7.19	8.39	3.2	2.95	2.2
(8)	0	100	100	100	100	0	99.97	0	0	0	7.47	8.44	3.29	2.89	2.12
(9)	0	100	99.94	100	100	0	99.92	0	99.96	0	8.19	8.42	3.48	2.91	1.99
(10)	0	100	100	100	100	0	100	0	0	0	7.19	8.37	3.2	2.95	2.2
(11)	0	100	100	100	100	0	100	0	100	0	7.44	8.4	3.29	2.89	2.11
(12)	0	100	100	100	100	0	100	100	100	100	8.09	100	3.48	2.83	1.98
(13)	31.67	60.69	81.47	63.23	98.85	11.52	6.38	62.36	82.59	6.31	9.55	18.75	24.47	38.01	34.62
(14)	60.84	15.03	89.47	14.61	97.11	22.47	67.26	34.51	83.81	54.52	8.51	5.8	30.61	41.19	14.93
(15)	69.52	82.52	65.26	77.03	74.3	87.84	12.96	99.4	72.4	17.96	3.72	32.99	16.82	32.58	8.66
(16)	84.84	88.71	63.42	81.72	93.6	69.72	12.62	8.34	11.95	19.53	13.49	38.13	8.75	27.16	22.45
(17)	10.27	89.76	70.9	86.48	99.91	55.89	59.07	16.85	73.04	8.57	18.38	18.22	19.83	10.17	3.62
(18)	100	0	0	100	100	0	0	0	100	0	0.03	0	0	0	0.26
(19)	0	0	100	100	100	100	0	0	0	0	0	0	0	0	0
(20)	0	100	0	100	100	0	0	100	0	0	0	0	0	0	0

(13)-(17)), since these are the only ones that distribute the resources subject to the "normal human thinking" so that no technical attribute level is restricted to zero.

Acknowledgements. The paper is supported by the European Union and co-financed by the European Regional Development Fund TAMOP-4.2.1/B-09/1/KONV-2010-0003: Mobility and Environment: Researches in the fields of motor vehicle industry, energetics and environment in the Middle- and West-Transdanubian Regions of Hungary.

References

[1] Balázs, K., Botzheim, J., Kóczy, L.T.: Comparative Investigation of Various Evolutionary and Memetic Algorithms. In: Rudas, I.J., Fodor, J., Kacprzyk, J. (eds.) Computational Intelligence in Engineering. SCI, vol. 313, pp. 129–140. Springer, Heidelberg (2010)

[2] Botzheim, J., Földesi, P.: Parametric representation of fuzzy power function for decision-making processes. In: Proceedings of the 7th International Symposium on Management Engineering, Kitakyushu, Japan, pp. 248–255 (2010)

[3] Botzheim, J., Cabrita, C., Kóczy, L.T., Ruano, A.E.: Fuzzy rule extraction by bacterial memetic algorithms. In: Proceedings of the 11th World Congress of International Fuzzy Systems Association, Beijing, China, pp. 1563–1568 (2005)

[4] Coello, C.A.C., Lamont, G.B.: Applications of Multi-Objective Evolutionary Algorithms. World Scientific (2004)
[5] Coello, C.A.C., Dhaenens, C., Jourdan, L.: Advances in Multi-Objective Nature Inspired Computing. Springer (2009)
[6] Fieldsend, J.E.: Multi-objective particle swarm optimisation methods. Department of Computer Science. University of Exeter (2004)
[7] Földesi, P., Botzheim, J.: Interpretation of Loss Aversion in Kano's Quality Model. In: Watada, J., Phillips-Wren, G., Jain, L.C., Howlett, R.J. (eds.) Intelligent Decision Technologies. SIST, vol. 10, pp. 165–174. Springer, Heidelberg (2011)
[8] Földesi, P., Kóczy, L.T., Botzheim, J.: Fuzzy solution for non-linear quality models. In: Proceedings of the 12th International Conference on Intelligent Engineering Systems, INES 2008, Miami, Florida, pp. 269–275 (2008)
[9] Hernandez-Diaz, A.G., Coello, C.A.C., Perez, F., Caballero, R., Molina, J., Santana-Quintero, L.V.: Seeding the initial population of a multi-objective evolutionary algorithm using gradient-based information. In: Proceedings of the IEEE Congress on Evolutionary Computation, Hong Kong, pp. 1617–1624 (2008)
[10] Kano, N., Seraku, N., Takahashi, F., Tsuji, S.: Attractive quality and must-be quality. The Journal of Japanese Society for Quality Control 14(2), 39–48 (1984)
[11] Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of the IEEE Int. Conf. on Neural Networks, Perth, Australia, pp. 1942–1948 (1995)
[12] Kennedy, J., Eberhart, R.C., Shi, Y.: Swarm Intelligence. Morgan Kaufmann (2001)
[13] Levenberg, K.: A method for the solution of certain non-linear problems in least squares. Quart. Appl. Math. 2(2), 164–168 (1944)
[14] Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math. 11(2), 431–441 (1963)
[15] Moscato, P.: On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms. Tech. Rep. Caltech, Pasadena, USA (1989)
[16] Mostaghim, S., Teich, J.: Strategies for finding good local guides in multi-objective particle swarm optimization (MOPSO). In: Proceedings of the IEEE Swarm Intelligence Symposium, pp. 26–33 (2003)
[17] Nawa, N.E., Furuhashi, T.: Fuzzy system parameters discovery by bacterial evolutionary algorithm. IEEE Transactions on Fuzzy Systems 7(5), 608–616 (1999)
[18] Shukla, P.K.: Exploiting Second Order Information in Computational Multi-objective Evolutionary Optimization. In: Neves, J., Santos, M.F., Machado, J.M. (eds.) EPIA 2007. LNCS (LNAI), vol. 4874, pp. 271–282. Springer, Heidelberg (2007)

Super Pairwise Comparison Matrix in the Multiple Dominant AHP

Takao Ohya and Eizo Kinoshita

Abstract. We have proposed a super pairwise comparison matrix (SPCM) to express all pairwise comparisons in the evaluation process of the dominant analytic hierarchy process (AHP) or the multiple dominant AHP (MDAHP) as a single pairwise comparison matrix. This paper shows, by means of a numerical counterexample, that in MDAHP an evaluation value resulting from the application of the logarithmic least-squares method (LLSM) to a SPCM does not necessarily coincide with that of the evaluation value resulting from the application of the geometric mean multiple dominant AHP (GMMDAHP) to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method.

Keywords: super pairwise comparison matrix, dominant AHP, the multiple dominant AHP, logarithmic least square method, the geometric mean multiple dominant AHP.

1 Introduction

The analytic hierarchy process (AHP) proposed by Saaty[1] enables objective decision making by top-down evaluation based on an overall aim.

In actual decision making, a decision maker often has a specific alternative (regulating alternative) in mind and makes an evaluation on the basis of the alternative. This was modeled in dominant AHP, proposed by Kinoshita and Nakanishi[2].

If there are more than one regulating alternatives and the importance of each criterion is inconsistent, the overall evaluation value may differ for each regulating alternative. As a method of integrating the importances in such cases, the concurrent convergence method (CCM) was proposed. Kinoshita and Sekitani[3] showed the convergence of CCM.

Takao Ohya
School of Science and Engineering, Kokushikan University, Tokyo, Japan

Eizo Kinoshita
Faculty of Urban Science, Meijo University, Gifu, Japan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 319–327.
springerlink.com

Meanwhile, Ohya and Kinoshita[4] proposed the geometric mean multiple dominant AHP (GMMDAHP), which integrates weights by using a geometric mean based on an error model to obtain an overall evaluation value. Herein, such methods of evaluation with multiple regulating alternatives will be generically referred to as the multiple dominant AHP (MDAHP).

Section 2 briefly explains dominant AHP and MDAHP and then proposes a super pairwise comparison matrix (SPCM) to express the pairwise comparisons appearing in the evaluation processes of the dominant AHP and MDAHP as a single pairwise comparison matrix.

Section 3 gives a specific numerical example. With the numerical example, it is shown that in MDAHP an evaluation value resulting from the application of the logarithmic least-squares method (LLSM) to a SPCM does not necessarily coincide with that of the evaluation value resulting from the application of the GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method.

2 MDAHP and SPCM

This section explains dominant AHP, GMMDAHP and a SPCM to express the pairwise comparisons appearing in the evaluation processes of dominant AHP and MDAHP as a single pairwise comparison matrix. Section 2.1 outlines dominant AHP procedure and explicitly states pairwise comparisons, and section 2.2 outlines GMMDAHP. Section 2.3 explains the SPCM that expresses these pairwise comparisons as a single pairwise comparison matrix.

2.1 Evaluation in Dominant AHP

The true absolute importance of alternative $a(a=1,...,A)$ at criterion $c(c=1,...,C)$ is v_{ca}. The final purpose of the AHP is to obtain the relative value (between alternatives) of the overall evaluation value $v_a = \sum_{c=1}^{C} v_{ca}$ of alternative a. The procedure of dominant AHP for obtaining an overall evaluation value is as follows:

Dominant AHP

Step 1: The relative importance $u_{ca} = \alpha_c v_{ca}$ (where α_c is a constant) of alternative a at criterion c is obtained by some kind of methods. In this paper, u_{ca} is obtained by applying the pairwise comparison method to alternatives at criterion c.

Step2: Alternative d is the regulating alternative. The importance u_{ca} of alternative a at criterion c is normalized by the importance u_{cd} of the regulating alternative d, and $u_{ca}^{d}\ (=u_{ca}/u_{cd})$ is calculated.

Step3: With the regulating alternative d as a representative alternative, the importance w_c^d of criterion c is obtained by applying the pairwise comparison method to criteria, where, w_c^d is normalized by $\sum_{c=1}^{C} w_c^d = 1$.

Step4: From u_{ca}^{d}, w_c^d obtained at Steps 2 and 3, the overall evaluation value $t_a = \sum_{c=1}^{C} w_c^d u_{ca}^d$ of alternative a is obtained. By normalization at Steps 2 and 3, $u_d = 1$. Therefore, the overall evaluation value of regulating alternative d is normalized to 1.

2.2 GMMDAHP

Let $\hat{w}_c^{d(d')}$ be the unknown evaluation value of the criterion c from the alternative $r \neq r'$. Then, Kinoshita and Nakanishi [2] propose a following evaluation rule under their assumption.

$$\hat{w}_c^{d(d')} = \frac{w_c^{d'} u_{dc} / u_{d'c}}{\sum_{c'} w_{c'}^{d'} u_{d'c'} / u_{d'c'}} \tag{1}$$

Let $\boldsymbol{D}$ be a set of regulating alternatives. MDAHP requires the decision maker to evaluate criteria from the viewpoint of each regulating alternative $d \in \boldsymbol{D}$. Let w_c^d be the evaluation value of the criterion c from the regulating alternative $d \in \boldsymbol{D}$. w_c^d are normalized as $\sum_c w_c^d = 1$ for all $d \in \boldsymbol{D}$. If evaluation values are consistent, all $\hat{w}_c^{d(d')}$ are same for all $d' \in D$. But, almost all evaluation involve inconsistency.

GMMDAHP calculate the weights for the criteria with the geometric means from multiple regulating alternatives as the following formula.

$$\hat{w}_c^d = \frac{\left(\prod_{d' \in D} w_c^{d'} u_{dc} / u_{d'c} \right)^{\frac{1}{|D|}}}{\sum_{c'} \left(\prod_{d' \in D} w_{c'}^{d'} u_{dc'} / u_{d'c'} \right)^{\frac{1}{|D|}}} \tag{5}$$

$\hat{w}_c^d$ is the geometric mean of $w_c^{d(d')}$ and multiplied constant of normalization as $\sum_c \hat{w}_c^d = 1$.

All $\hat{w}_c^d$ always have following relations.

$$\hat{w}_c^d = \frac{\hat{w}_c^{d'} u_{dc} / u_{d'c}}{\sum_{c'} \hat{w}_{c'}^{d'} u_{dc'} / u_{d'c'}} \tag{6}$$

So, all $\hat{w}_c^d$ are consistent.

2.3 SPCM

The relative comparison values $r_{c'a'}^{ca}$ of importance v_{ca} of alternative a at criteria c as compared with the importance $v_{c'a'}$ of alternative a' in criterion c', are arranged in a (CA × CA) or (AC × AC) matrix. In a (CA × CA) matrix, index of alternative changes first. In a (AC × AC) matrix, index of criteria changes first. This is proposed as the SPCM $R = (r_{c'a'}^{ca})\ or\ (r_{a'c'}^{ac})$.

In a SPCM, symmetric components have a reciprocal relationship as in pairwise comparison matrices. Diagonal elements are 1 and the following relationships are true:

$$r_{ca}^{c'a'} = 1 \Big/ r_{c'a'}^{ca}, \tag{7}$$

$$r_{ca}^{ca} = 1. \tag{8}$$

Pairwise comparison at Step 1 of dominant AHP consists of the relative comparison value $r_{ca'}^{ca}$ of importance v_{ca} of alternative a, compared with the importance $v_{ca'}$ of alternative a' at criterion c.

Pairwise comparison at Step 3 of dominant AHP consists of the relative comparison value $r_{c'd}^{cd}$ of importance v_{cd} of alternative d at criterion c, compared with the importance $v_{c'd}$ of alternative d at criterion c', where the regulating alternative is d.

Fig. 1 and 2 show SPCMs using the MDAHP when there are three alternatives (1-3) and four criteria (I-IV) and all alternatives are regulating ones. In these figures, * represents pairwise comparison between alternatives in a same criterion and # represents pairwise comparison between criteria of a same alternative.

	I 1	I 2	I 3	II 1	II 2	II 3	III 1	III 2	III 3	IV 1	IV 2	IV 3
I 1	1	*	*	#			#			#		
I 2	*	1	*		#			#			#	
I 3	*	*	1			#			#			#
II 1	#			1	*	*	#			#		
II 2		#		*	1	*		#			#	
II 3			#	*	*	1			#			#
III 1	#			#			1	*	*	#		
III 2		#			#		*	1	*		#	
III 3			#			#	*	*	1			#
IV 1	#			#			#			1	*	*
IV 2		#			#			#		*	1	*
IV 3			#			#			#	*	*	1

Fig. 1 SPCM by MDAHP (CA × CA)

	I 1	II 1	III 1	IV 1	I 2	II 2	III 2	IV 2	I 3	II 3	III 3	IV 3
I 1	1	#	#	#	*				*			
II 1	#	1	#	#		*				*		
III 1	#	#	1	#			*				*	
IV 1	#	#	#	1				*				*
I 2	*				1	#	#	#	*			
II 2		*			#	1	#	#		*		
III 2			*		#	#	1	#			*	
IV 2				*	#	#	#	1				*
I 3	*				*				1	#	#	#
II 3		*				*			#	1	#	#
III 3			*				*		#	#	1	#
IV 3				*				*	#	#	#	1

Fig. 2 SPCM by MDAHP (AC × AC)

SPCM of dominant AHP or MDAHP is an incomplete pairwise comparison matrix. Therefore, the LLSM based on an error model or an eigenvalue method such as the Harker[6] method or two-stage method is applicable to the calculation of evaluation values from an SPCM.

3 Numerical Example

Three alternatives from 1 to 3 and four criteria from I to IV are assumed, where Alternative 1 is the regulating alternative.

As the result of pairwise comparison between alternatives at criteria c (c = I,...,IV), the following pairwise comparison matrices $R_c^A, c = \text{I},...,\text{IV}$ are obtained:

$$R_{\text{I}}^A = \begin{pmatrix} 1 & 1/3 & 5 \\ 3 & 1 & 3 \\ 1/5 & 1/3 & 1 \end{pmatrix}, \quad R_{\text{II}}^A = \begin{pmatrix} 1 & 7 & 3 \\ 1/7 & 1 & 1/3 \\ 1/3 & 3 & 1 \end{pmatrix},$$

$$R_{\text{III}}^A = \begin{pmatrix} 1 & 1/3 & 1/3 \\ 3 & 1 & 1/3 \\ 3 & 3 & 1 \end{pmatrix}, \quad R_{\text{IV}}^A = \begin{pmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 1 \\ 1/5 & 1 & 1 \end{pmatrix}$$

With regulating alternatives 1 to 3 as the representative alternatives, importance between criteria was evaluated by pairwise comparison. As a result, the following pairwise comparison matrix $R_{\text{I}}^C, R_2^C, R_3^C$ is obtained:

$$R_1^C = \begin{pmatrix} 1 & 1/3 & 3 & 1/3 \\ 3 & 1 & 3 & 1 \\ 1/3 & 1/3 & 1 & 1/3 \\ 3 & 1 & 3 & 1 \end{pmatrix}, \quad R_2^C = \begin{pmatrix} 1 & 9 & 1 & 5 \\ 1/9 & 1 & 1/3 & 1 \\ 1 & 3 & 1 & 1 \\ 1/5 & 1 & 1 & 1 \end{pmatrix},$$

$$R_3^C = \begin{pmatrix} 1 & 1/3 & 1/9 & 3 \\ 3 & 1 & 1/5 & 5 \\ 9 & 5 & 1 & 5 \\ 1/3 & 1/5 & 1/5 & 1 \end{pmatrix}$$

Fig. 3 shows the SPCM for this example.

(1) SPCM +LLSM

Table 1 shows the evaluation values obtained from the SPCM in Fig. 3.

$$\begin{pmatrix}
1 & 1/3 & 5 & 1/3 & & & 3 & & & 1/3 & & \\
3 & 1 & 3 & & 9 & & & 1 & & & 5 & \\
1/5 & 1/3 & 1 & & & 1/3 & & & 1/9 & & & 3 \\
3 & & & 1 & 7 & 3 & 3 & & & 1 & & \\
 & 1/9 & & 1/7 & 1 & 1/3 & & 1/3 & & & 1 & \\
 & & 3 & 1/3 & 3 & 1 & & & 1/5 & & & 5 \\
1/3 & & & 1/3 & & & 1 & 1/3 & 1/3 & 1/3 & & \\
 & 1 & & & 3 & & 3 & 1 & 1/3 & & 1 & \\
 & & 9 & & & 5 & 3 & 3 & 1 & & & 5 \\
3 & & & 1 & & & 3 & & & 1 & 3 & 5 \\
 & 1/5 & & & 1 & & & 1 & & 1/3 & 1 & 1 \\
 & & 1/3 & & & 1/5 & & & 1/5 & 1/5 & 1 & 1
\end{pmatrix}$$

Fig. 3 SPCM

Table 1 Evaluation values obtained by SPCM+LLSM

	Criterion I	Criterion II	Criterion III	Criterion IV	Overall evaluation value
Alternative 1	0.1758	0.3970	0.0961	0.3312	1
Alternative 2	0.3386	0.0570	0.1851	0.0838	0.6644
Alternative 3	0.0707	0.1636	0.4655	0.0525	0.7522

(2) GMMDAHP + geometric mean

Table 2 shows the evaluation values resulting from the application of the GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method..

Table 2 Evaluation values obtained by GMMDAHP and geometric mean

	Criterion I	Criterion II	Criterion III	Criterion IV	Overall evaluation value
Alternative 1	0.1923	0.4289	0.0977	0.2810	1
Alternative 2	0.3375	0.0564	0.2033	0.0790	0.6761
Alternative 3	0.0658	0.1555	0.4229	0.0666	0.7108

The evaluation value resulting from the application of LLSM to a SPCM shown in Table 2 does not necessarily coincide with that of the evaluation value resulting from the application of the GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method shown in Table 2.

With these numerical results, it is shown that the evaluation value resulting from the application of LLSM to the SPCM does not necessarily coincide with the evaluation value resulting from the application of GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method.

(3) SPCM + Harker method

In the Harker method, the value of a diagonal element is set to the number of missing entries in the row plus 1 and then evaluation values are obtained by the usual eigenvalue method.

Fig. 4 shows the SPCM by the Harker method. Table 3 shows the evaluation values obtained from the SPCM in Fig. 4.

$$\begin{pmatrix}
7 & 1/3 & 5 & 1/3 & & & 3 & & & 1/3 & & \\
3 & 7 & 3 & & 9 & & & 1 & & & 5 & \\
1/5 & 1/3 & 7 & & & 1/3 & & & 1/9 & & & 3 \\
3 & & & 7 & 7 & 3 & 3 & & & 1 & & \\
 & 1/9 & & 1/7 & 7 & 1/3 & & 1/3 & & & 1 & \\
 & & 3 & 1/3 & 3 & 7 & & & 1/5 & & & 5 \\
1/3 & & & 1/3 & & & 7 & 1/3 & 1/3 & 1/3 & & \\
 & 1 & & & 3 & & 3 & 7 & 1/3 & & 1 & \\
 & & 9 & & & 5 & 3 & 3 & 7 & & & 5 \\
3 & & & 1 & & & 3 & & & 7 & 3 & 5 \\
 & 1/5 & & & 1 & & & 1 & & 1/3 & 7 & 1 \\
 & & 1/3 & & & 1/5 & & & 1/5 & 1/5 & 1 & 7
\end{pmatrix}$$

Fig. 4 SPCM by the Harker method

Table 3 Evaluation values obtained by the Harker method to SPCM

	Criterion I	Criterion II	Criterion III	Criterion IV	Overall evaluation value
Alternative 1	0.1885	0.3986	0.0977	0.3152	1
Alternative 2	0.3444	0.0551	0.2002	0.0858	0.6856
Alternative 3	0.0742	0.1609	0.4550	0.0611	0.7512

(4) GMMDAHP＋the eigenvalue method

Table 4 shows the evaluation values resulting from the application of the GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the eigenvalue method.

Table 4 Evaluation values obtained by GMMDAHP and the eigenvalue method

	Criterion I	Criterion II	Criterion III	Criterion IV	Overall evaluation value
Alternative 1	0.1936	0.4182	0.0996	0.2885	1
Alternative 2	0.3397	0.0549	0.2072	0.0811	0.6830
Alternative 3	0.0662	0.1516	0.4311	0.0684	0.7173

4 Conclusion

Ohya and Kinoshita[5] shows that the evaluation value resulting from the application of LLSM to an SPCM agrees with the evaluation value determined by the application of the dominant AHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean.

In contrast with that, this paper shows, by means of a numerical counterexample, that in MDAHP an evaluation value resulting from the application of LLSM to a SPCM does not necessarily coincide with that of the evaluation value resulting from the application of GMMDAHP to the evaluation value obtained from each pairwise comparison matrix by using the geometric mean method.

References

1. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
2. Kinoshita, E., Nakanishi, M.: Proposal of new AHP model in light of dominative relationship among alternatives. Journal of the Operations Research Society of Japan 42, 180–198 (1999)
3. Kinoshita, E., Sekitani, K., Shi, J.: Mathematical Properties of Dominant AHP and Concurrent Convergence Method. Journal of the Operations Research Society of Japan 45, 198–213 (2002)
4. Ohya, T., Kinoshita, E.: The Geometric Mean Concurrent Convergence Method. In: Proceedings of the 10th International Symposium on the Analytic Hierarchy Process (2009) ISSN 1556-8296
5. Ohya, T., Kinoshita, E.: Proposal of Super Pairwise Comparison Matrix. In: Watada, J., Phillips-Wren, G., Jain, L.C., Howlett, R.J., et al. (eds.) Intelligent Decision Technologies. SIST, vol. 10, pp. 247–254. Springer, Heidelberg (2011)
6. Harker, P.T.: Incomplete pairwise comparisons in the Analytic Hierarchy Process. Mathematical Modeling 9, 837–848 (1987)

Visualization of Rules in Rule-Based Classifiers

Susanne Bornelöv, Stefan Enroth, and Jan Komorowski

Abstract. Interpretation and visualization of the classification models are important parts of machine learning. Rule-based classifiers often contain too many rules to be easily interpreted by humans, and methods for post-classification analysis of the rules are needed. Here we present a strategy for circular visualization of sets of classification rules. The Circos software was used to generate graphs showing all pairs of conditions that were present in the rules as edges inside a circle. We showed using simulated data that all two-way interactions in the data were found by the classifier and displayed in the graph, although the single attributes were constructed to have no correlation to the decision class. For all examples we used rules trained using the rough set theory, but the visualization would by applicable to any sort of classification rules. This method for rule visualization may be useful for applications where interaction terms are expected, and the size of the model limits the interpretability.

1 Introduction

An important part of machine learning and classification is the visualization and interpretability of the model (Thearling et al 2001). Certain classifiers, e.g. neural networks and support vector machines, transform the original attributes into combinations or functions thereof. Other classifiers, e.g. rule-based classifiers and decision trees keep the original attributes, which allow non-experts to interpret the

Susanne Bornelöv · Jan Komorowski
Department of Cell and Molecular Biology, Science for Life Laboratory,
Biomedical Center, Uppsala University, Box 598, SE-75124 Uppsala, Sweden

Stefan Enroth,
Department of Immunology, Genetics and Pathology, Science for Life Laboratory,
Rudbeck Laboratory, Uppsala University, Uppsala, Sweden

Jan Komorowski,
Interdisciplinary Centre for Mathematical and Computational Modelling,
University of Warsaw, Warszawa, Poland

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 329–338.
springerlink.com

model after only a short introduction. We have been working with rule-based classifiers, which have the advantage of a very simple representation. A single rule has an antecedent (the ”IF-part”) and a consequent part (the ”THEN-part”). The rule antecedent contains a conjunction of any finite number of conditions which, if fulfilled, predict a specific decision class in the consequent. Each condition in the antecedent consists of an attribute and an attribute value. An example of a rule is: IF $Attribute_0$=$value_0$ AND $Attribute_1$=$value_1$ THEN $Class$=C_0.

Note that these rules are trained from tabular data, in which the objects are the rows and the attributes are the columns. The attributes are divided into condition attributes and the class attribute. The aim of the classification is to predict the decision class using the information from the condition attributes. A typical application could be to predict the cancer site of origin based on gene expression data, in which the patients would be the objects, the gene expression the condition attributes, and the site of origin the decision class.

There are different algorithms for rule generation, including rough set-based classification rules and association or classification rule mining. Association rules (Agrawal et al 1993) differ from the rough set-based classification rules, but are interesting from a rough set perspective as a lot of effort has been put into the visualization of them. In association rules certain observations are linked together to form a rule and both the antecedent and the consequent of the rule contain any set (called an itemset) of such observations. A common problem for both classification and association rules is that thousands or tens of thousands of such rules can be constructed by the algorithms resulting in a model that is hard to interpret because of the high number of rules.

Different filtering and rule generalization algorithms can be applied to reduce the number of rules, but could possibly harm the model by oversimplification. Many rules could indeed be needed to describe a complicated system and a general method for visualization of the rules would thus be of high usability.

Interpretation of rough rules has previously been discussed (Greco et al 2002). In addition there are many overviews of visualization techniques for association rule mining (Buono 2005, Bruzzese and Davino 2008). The simplest technique is the tabular visualization, in which each row is a rule and each column is an itemset. Similar methods have been used for rough set-based rules (Kontijevskis et al 2007, Kierczak et al 2009) in which each column represented an attribute. This method has the disadvantage that the user only can see a handful of rules at the same time. A two-dimensional matrix is similar but mostly suited where the itemsets contain only one observation, with the antecedent on the x-axis and the consequent on the y-axis. Additionally, similar itemsets can be clustered together to make it more comprehensible and information can be encoded by colors and sizes in the matrix cells, as implemented in the R package arulesViz (Hahsler and Chelluboina 2011). Another technique that focuses primarly on the rule statistics is the Two-key plot in which the rules are described by their statistics and pairwise relations, but not based on the items in the rule. Different graph-based methods are more oriented to showing the itemsets in the rules. One such interesting approach is a circular graph in which items in the rules are mapped around a circle and

edges are connecting them according to the associations (Rainsford and Roddick 2000). Other types of graphs will e.g. connect the antecedent of the rule with the consequent by an arrow, and use colors or labels to encode rule statistics, but become hard to read when there are many rules. Various more advanced 3D visualization techniques have been developed as well, but do not scale well to applications where many rules are generated.

Draminiski et al. used another approach and created a network, not based on the classification model, but on the preceding feature selection as a part of the Monte Carlo feature selection interdependency discovery algorithm (Draminiski et al 2010). During the feature selection, they created a large number of subsets of the original data (both in the attribute and object space) and used those to train classification trees. Pairs of attributes got scores when they occurred closely together in those trees, and the scores were shown as a network.

Traditional networks are hard to compare, as the positions of the nodes and edges are highly dependent on implementation details instead of properties of the data. The Circos software for visualizing data and information in a circular layout (Krzywinski et al 2009) has provided a useful tool to solve this problem. Every node in the circular graphs can be placed on a circle in an order that is dependent of characteristics of the data itself. Similar nodes will then have similar positions in similar graphs, which make the graphs easier to compare and interpret.

We have developed a method to visualize classification rules using a circular graph layout with some similarities to the one previously used for association rules. Colors and saturation are used to represent the rule statistics (support and accuracy) and we have been able to include the attribute values and the decision classes in order to more precisely describe which class is described by a specific conditions. The method assumes that pairs of conditions that often occur in the rules are those that are most interesting and also that the pairs are more important if the rules they occur in are of high quality. The graphs were created using the Circos software to produce high-quality highly aesthetic graphs.

Here we present examples of our approach for two different classification problems, based both on real and simulated data, and discuss the interpretation of the graphs. We believe that this new method for displaying the rule-based classifiers will increase the usability of rule-based classifiers..

2 Rule Training and Visualization

In this section we first give additional terminology of the rules and then we present how the rules and the circular visualization were generated.

2.1 Rule Statistics

The rule *support* of a rule was defined as the number of objects in the data set that fulfill the antecedent or the left-hand side (LHS) of the rule. The rule *accuracy* was defined as the fraction of those objects that also fulfilled the consequent or the

right-hand side (RHS) of the rule or in other words the LHS-conditioned probability of the RHS:

$$accuracy = P(RHS|LHS). \quad (1)$$

If a rule had support 50 and accuracy 0.8 it would then tell us that the rule was based on 50 examples that fulfilled the antecedent of the rule, and that 0.8*50=40 of those examples were correctly classified by the rule.

2.2 Training of the Classifiers

The circular visualization of the rules is not limited to any specific method or algorithm for training of the classifier, but it will require the classifier to contain only a very limited set of consequents. We used the Rosetta system (Komorowski et al 2002) based on rough set theory to generate the rule-based classifiers. For large data sets, a feature selection was performed prior to the training of the classifier to remove less important attributes and thus increase the model stability and reduce the computational need. The accuracy of the model was normally estimated using cross validation.

The process of training the classifiers contains the following steps:

1. Prepare the data into a tabular form with objects as rows and attributes as columns. The last column should be the decision class.
2. (*optionally*) Perform feature selection to reduce the number of attributes.
3. (*optionally, if applicable*) Select a balanced subset of the dataset with respect to the value of the decision class.
4. Train a rule based classifier using Rosetta:
   - (*if applicable*) Discretize numeric attributes into intervals.
   - (*if applicable*) Remove/replace missing values.
   - Compute reducts and generate rules.
   - (*optionally*) Filter the rules according to rule quality.
   - (*optionally*) Simplify the rules by a rule tuning algorithm.
5. Perform a cross validation using the same settings to assess the quality of the classifier.

An uneven class distribution could cause difficulties. Optionally, for such data sets the steps 4-5 have been repeated for multiple subsets balanced by undersampling. The rules from all runs have then been merged into one rule set, and the rule statistics has been re-calculated using the whole training set.

2.3 Translation of the Rules Into Pairwise Scores

A classifier consists of multiple rules for different decision classes. We divided the rules based on the decision class. If there were multiple classes in the consequent, we considered only the largest. One graph was constructed for each decision class.

Each condition that appeared in the rules was represented as one node in the graph. Assuming that three attributes (A_1, A_2, A_3) with two possible values (v_1, v_2) each were present in the dataset, this would give the following 2*3 nodes: 'A_1=v_1', 'A_1=v_2', 'A_2=v_1', 'A_2=v_2', 'A_3=v_1', and 'A_3=v_2'. A pair of nodes (conditions) was connected in the graph if they appeared in the same rule(s).

Let $R(x,y)$ be the set of all rules that contained two nodes x and y. The connection strength between x and y was defined as

$$\text{connection}(x, y) = \sum_{r \in R(x,y)} \text{support}(r) \cdot \text{accuracy}(r) \tag{2}$$

which will represent the sum of all objects that were correctly classified by rules containing a certain pair of conditions.

2.4 Graph Construction

We translated the pairwise scores into a table. As the connections are not directed, each score was only printed once in the table. The Circos tool *tableviewer* was run with this table as input to produce the circular graph. Additional parameters were changed in the Circos configuration files to fine-tune the behavior and colors of the circle.

3 Interpretation of the Circular Graphs

In this section we describe the rule visualization and give two examples of circular graphs created for different classification problems.

3.1 Description of the Circular Graph

The combinations of attributes in the rules are visualized as a circular graph. Each condition that appears in the rules will be positioned on the circle. Ideally, the conditions are positioned in an orderly manner (following alphabetical order or by a problem-dependent property, e.g. by genomic position, or residue position in a protein). The nodes can be colored either individually, or by an arbitrary group. A thinner line to the inside of the nodes shows the color of the node on the other side of the edge. The edges are positioned inside the circle, and their color and width is dependent on the number of objects in the dataset that were correctly classified by rules containing the specific combination of conditions (*support* * *accuracy*). Not all possible pairs of conditions are present in rules, and thus there are many pairs which are not connected in the graph. The upper quartile of the edges is colored as a scale from yellow (lowest score) to red (highest score), and the three lower quartiles are colored in light gray. The width of the nodes (conditions) on the circle is the sum of the width of all edges that go to that node. The total width of all edges is automatically scaled by Circos to fit the whole circle.

A limitation of the graphs is that only rules with at least two conditions in the antecedent of the rule can be included in the graph. From the other point, the view

of rules sets which contain mostly rules longer than two conditions might be oversimplified in the circular view. The best application would be where the rules could be expected to contain many two-way interactions or where the interactions are of specific interest. If it is thought to be an issue, the number of conditions in the rules may be easily counted prior to graph generation.

3.2 Second Order Interactions Using Simulated Data

We simulated data in order to investigate how pair-wise correlations were displayed in the graphs. We created a dataset with 20 binary attributes and one binary decision class. The attributes were divided in two groups, one with R_1, R_2, … R_{10} and the other with S_1, S_2, …, S_{10}. For each pair of R_i, S_i ($1 \leq i \leq 10$), the values for the *R* attribute were chosen at random with equal probabilities for 0 and 1, and the corresponding *S* attribute values were similarly chosen at random with 50 % probability, and 50 % chance to follow the definition: 1 if the value of R_i corresponded to the class label, and 0 otherwise.

Using this schema, we constructed a dataset with 20 attributes that were all uncorrelated to the decision attribute, but where a correlation between ten distinct pairs of attributes (R_1 and S_1, R_2 and S_2, …, R_{10} and S_{10}) and the decision class was expected. We wanted to observe two things: whether the pairwise correlations were described by the rules, and if they were visible in the circular graphs.

Rules were generated using the whole data set and the Rosetta software. We used an algorithm called *GeneticReducer* that normally produces many rules to simulate the difficulty of interpretability. A total of 91401 rules were returned from Rosetta. We used a 10-fold cross validation to measure the performance of the classifier and an average accuracy of 94 % was observed for the classifier, indicating good overall performance. Other parameters and algorithms for reduct calculation were tested and gave similar performance, but fewer rules (results not shown). A circular visualization of the rules was created (Fig. 1). Despite the huge number of rules and the low quality for some of the rules due to the the lack of rule filtering, the second order interactions were clearly found to be dominating the connections in the graph. For example, for the decision class 0 (Fig. 1, left), the strongest connection that goes from S4-0 (S4=0) goes to R4-1 (R4=1). The same pattern can be seen for all of the attributes, for both of the decision classes (Fig. 2). As the "*S*" attributes are markers of whether the "*R*" attributes correspond to the class label or not, the class should be 0 if they are of different value and 1 if they are of the same value (Table 1). These combinations are also present and possible to identify in the circular graph.

Table 1 Class dependence on the *R* and *S* attributes.

R	*S*	*Class*	Interpretation
0	0	1	*R* is 0, and the *S* attribute indicate that the class is not 0
0	1	0	*R* is 0, and the *S* attribute indicate that the class is 0
1	0	0	*R* is 1, and the *S* attribute indicate that the class is not 1
1	1	1	*R* is 1, and the *S* attribute indicate that the class is 1

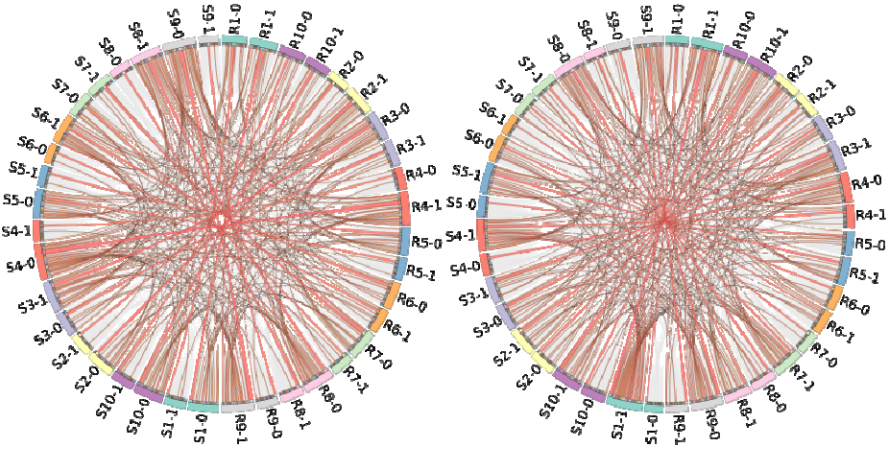


Fig. 1 A circular visualization for simulated data with two decision classes (0 to the left and 1 to the right) based on 91401 rules. The data was constructed so that no attribute were correlated to the decision, but all pairs of attributes R_i, S_i are correlated. Pairs of R_i and S_i terms were colored using the same color and are positioned on opposite sides of the circle centre. The connections between the R_i, S_i pairs are the dominating connections in the graph and for almost all nodes this is the strongest connection. A high resolution copy of this figure is available at http://www.anst.uu.se/susbo258/rulevis/Fig1.svg.

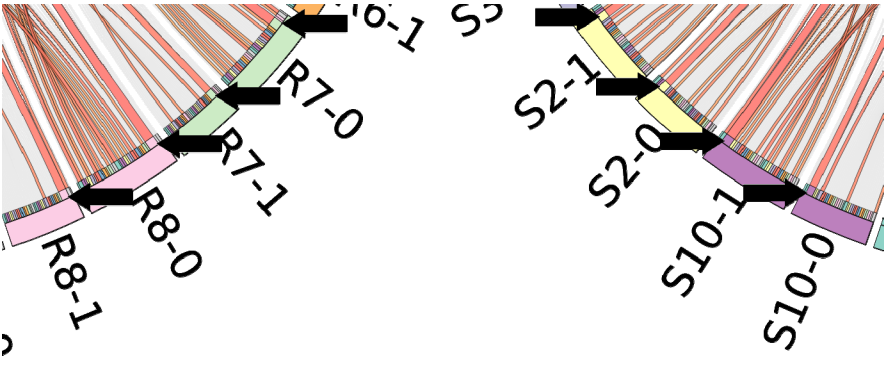


Fig. 2 A subpart of Fig. 1. For every node the strongest connection goes to a corresponding node of equivalent color and number on the other side of the circle, illustrated by the inner band where the widest segment (marked by an arrow) has the same color as the node itself.

3.2.1 Tabular Visualization of the Rules

As an alternative to showing the rules in a circular graph, the complete rules can be displayed using a tabular view. The rules for the simulated data in Section 3.2 were sorted by support, and the top 15 rules for decision class 0 were compiled into a table (Table 2). The table is very good for a detailed look at the best rules, but can only show a small fraction of all 91401 rules, of which 41785 were for

class 0. As expected (shown in Table 1), the best rules for class 0 consist of pairs of attributes R_i and S_i with different attribute values.

Table 2 Rules for the simulated data set presented in tabular form. Only the 15 rules with highest support for decision class 0 are shown. Each row represent one rule, and the columns R_1, S_1; R_2, S_2; …, R_{10}, S_{10} show which attributes that are part of the rule, and which value they are required to have (0 or 1). If there is only a dash in the cell, if means that the rule is independent of that particular attribute. Accuracy and support show rule statistics.

#	R1	S1	R2	S2	R3	S3	R4	S4	R5	S5	R6	S6	R7	S7	R8	S8	R9	S9	R10	S10	Accuracy	Support
1	-	-	-	-	-	-	-	-	1	0	-	-	-	-	-	-	-	-	-	-	0.715	274
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	1	0.724	268
3	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.758	260
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	1	-	-	-	-	0.771	253
5	-	-	-	-	-	-	-	-	-	-	-	-	1	0	-	-	-	-	-	-	0.723	249
6	-	-	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.758	248
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0	-	-	0.798	247
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0	0.751	245
9	-	-	-	-	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.739	245
10	-	-	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.722	245
11	-	-	-	-	-	-	1	0	-	-	-	-	-	-	-	-	-	-	-	-	0.803	244
12	-	-	-	-	-	-	0	1	-	-	-	-	-	-	-	-	-	-	-	-	0.713	244
13	-	-	-	-	-	-	-	-	-	-	0	1	-	-	-	-	-	-	-	-	0.756	242
14	-	-	-	-	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.786	238
15	-	-	-	-	-	-	-	-	-	-	-	-	0	1	-	-	-	-	-	-	0.756	238

3.3 Applications

We have together with collaborators applied this method for circular visualization of rules to a variety of problems, including epigenetics, the Alzheimer's disease, avian influenza, and allergy and asthma in children. Here we describe the epigenetic application (Enroth et al 2012) more in detail.

The aim of the classification was to predict the exon inclusion level based on post-translational histone modifications measured by Chromatin Immunoprecipitation followed by parallel sequencing (ChIP-seq). For each exon and histone modification the signal was discretized into a presence/absence representation for three different positions relative the exon (preceding, on, or succeeding). The rule model contained 165 rules after filtering, and the circular visualization (Fig. 3) made it easier to interpret the results. In concordance with the current literature, the most frequent attribute in the rules for the spliced out exons was the H3K36me3 histone modification absent from the intron succeeding the exon, represented as 'B S0'. The strongest connections for this attribute for the

'spliced out' decision class was 'J P1' (H4K91ac present preceding the exon), 'I S1' (H4K20me1 present succeeding the exon), and 'I P1' (H4K20me1 present preceding the exon). This sort of visualization could give new information about how the histone modifications act together, and it would have been hard to identify the same patterns only by looking at the rules.

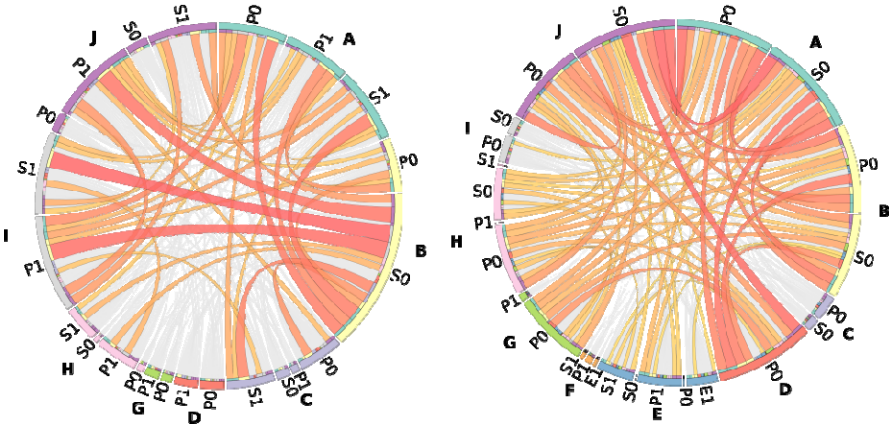


Fig. 3 Circular visualizations of the rules for spliced out exons (left) and included exons (right). The letters A-J encode the different histone modifications signals (A=H2BK5me1, B=H3K36me3, C=H3K4me1, D=H3K9me1, E=H3K9me2, F=H3K9me3, G=H3R2me1, H=H4K16ac, I=H4K20me1, J=H4K91ac), the S/E/P shows the position relative to the exons (S=succeeding, E=on the exon, P=preceding), and the number indicate if the modification should be absent (0) or present (1) in the rule. The figure was previously published in (Enroth et al 2012).

4 Conclusions

Visualization and interpretation of the results are both important parts of all machine learning. We have defined a circular visualization of pairs of conditions in rule-based classifiers. Our graphs are two-dimensional and easy to interpret, but still based both on the antecedent, consequent and the rule statistics (support and accuracy). Although support and accuracy are not indiscernible in the approach we suggest, separate graphs can, if needed, easily be constructed based solely on one of them. The method scales very well with the number of rules and is suitable as soon as the model size makes the interpretation hard. The visualization will however only show combinations of attributes, and is not suited if only first-order correlations are expected to be found.

It is important to note that the visualization can only show what is present in the rules. All algorithms for rule generation are normally heuristic. That means that not all combinations are guaranteed to be found, and those that are found are not necessarily the best ones. There is also a trade-off between overfitting and correctness. Taking the average over multiple runs will reduce the stochastic variation, but cannot compensate for algorithmic biases.

References

Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules Between Sets of Items in Large Databases. In: SIGMOD Conference, pp. 207–216 (1993)

Bruzzese, D., Davino, C.: Visual Mining of Association Rules. In: Simoff, S.J., Böhlen, M.H., Mazeika, A. (eds.) Visual Data Mining. LNCS, vol. 4404, pp. 103–122. Springer, Heidelberg (2008), doi:10.1007/978-3-540-71080-6_8

Buono, P., Costabile, M.F.: Visualizing Association Rules in a Framework for Visual Data Mining. In: Hemmje, M., Niederée, C., Risse, T. (eds.) From Integrated Publication and Information Systems to Information and Knowledge Environments. LNCS, vol. 3379, pp. 221–231. Springer, Heidelberg (2005), doi:10.1007/978-3-540-31842-2_22

Dramiński, M., Kierczak, M., Koronacki, J., Komorowski, J.: Monte Carlo Feature Selection and Interdependency Discovery in Supervised Classification. In: Koronacki, J., Raś, Z.W., Wierzchoń, S.T., Kacprzyk, J. (eds.) Advances in Machine Learning II. SCI, vol. 263, pp. 371–385. Springer, Heidelberg (2010)

Enroth, S., Bornelöv, S., Wadelius, C., Komorowski, J.: Combinations of histone modifications mark exon inclusion levels. PLoS ONE 7(1), e29911 (2012), doi:10.1371/journal.pone.0029911

Greco, S., Pawlak, Z., Słowiński, R.: Generalized Decision Algorithms, Rough Inference Rules, and Flow Graphs. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 93–104. Springer, Heidelberg (2002), doi:10.1007/3-540-45813-1_12

Hahsler, M., Chelluboina, S.: Visualizing Association Rules in Hierarchical Groups. Interface, 1–11 (2011)

Kierczak, M., Ginalski, K., Dramiński, M., Koronacki, J., Rudnicki, W.R., Komorowski, J.: A Rough Set Model of HIV-1 Reverse Transcriptase Resistome. Bioinformatics and Biology Insights 3, 109–127 (2009)

Komorowski, J., Øhrn, A., Skowron, A.: The Rosetta Rough Set Software System. In: Klösgen, W., Zytkow, J. (eds.) Handbook of Data Mining and Knowledge Discovery. Oxford University Press (2002)

Kontijevskis, A., Wikberg, J., Komorowski, J.: Computational Proteomics Analysis of HIV-1 Protease Interactome. Proteins 68, 305–312 (2007), doi:10.1002/prot.21415

Krzywinski, M.I., Schein, J.E., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra, M.A.: Circos: an Information Aesthetic for Comparative Genomics. Genome. Res. 19, 1639–1645 (2009), doi:10.1101/gr.092759.109

Rainsford, C.P., Roddick, J.: Visualisation of Temporal Interval Association Rules. In: Leung, K.-S., Chan, L., Meng, H. (eds.) IDEAL 2000. LNCS, vol. 1983, pp. 91–96. Springer, Heidelberg (2000)

Thearling, K., Becker, B., De Coste, D., Mawby, B., Pilote, M., Sommerfield, D.: Visualizing Data Mining Models. In: Fayyad, U., Grinstein, G., Wierse, A. (eds.) Information Visualization in Data Mining and Knowledge Discovery, pp. 205–222. Morgan Kaufmann, San Francisco (2001)

Part II
Applications of Intelligent Decision Technology

A Dynamic Programming Decision-Making Model of Investment Policy for Unconventional Oil Exploitation

Jinfeng Sun and Xiangpei Hu

Abstract. In this paper we focus on investment policy problems of unconventional oil exploitation and set up *a dynamic programming model* to help decision makers decide how to allocate the limited resources among a set of unconventional oil projects to maximize the total expected profits from investment horizon. Firstly, the urgency and feasibility of developing unconventional oil were introduced. Secondly, the properties of unconventional oil resources and the difficulty and complexity of exploiting them were analyzed. Thirdly, a multi-stage decision model was developed to help oil companies select an optimal investment policy given return on investment. Finally, a numerical example was provided by applying the model through backward recursion algorithm. The results demonstrate that the dynamic programming model provides an effective and efficient decision support tool for optimal investment policy of unconventional oil exploitation.

Keywords: dynamic programming model, unconventional oil exploitation, investment policy, backward recursion algorithm.

1 Introduction

Unconventional oil production such as Canadian oil sands production, Venezuelan extra-heavy oil production, and US shale oil production may play an extremely significant strategic role for most countries in the world to bridge the coming gap between the soaring oil demand and the declining oil supply although the oil peak

Jinfeng Sun · Xiangpei Hu
Institute of Systems Engineering, Dalian University of Technology, Dalian, 116024, P.R. China
e-mail: sunjinfeng123@126.com, drhxp@dlut.edu.cn

Jinfeng Sun
School of Economics and Management, China University of Petroleum (Huadong), Qingdao, 266555, P.R. China

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 341–351.
springerlink.com

time is far from agreement. Nevertheless, unconventional oil production is much more difficult, complex, and costly than conventional oil production such that oil companies have to spend many years and billions of dollars in exploring and exploiting unconventional oil as well as dealing with environment problems. Oil companies will decide whether investing unconventional oil or not, which one will be exploited, and how many resources should be allocated to get maximum return from investment planning horizon. Oil companies will suffer a direct economic loss except for the huge sunk costs as well as environment damages once decision at any decision-making process is unreasonable. The quality of decision this year has a direct impact on the profits and investment policy next year even in the life-cycle. Hence, it is absolutely crucial for oil companies to take a rigorous method to assess and improve the whole decision-making process for unconventional oil exploitation. The model developed here provides an effective and efficient tool to help decision makers to apportion a fixed amount of resources among a set of unconventional oil projects with different return on investment (ROI) to maximize the total expected profits from investment horizon.

The objective of this study is to optimize investment policy for unconventional oil exploitation. A dynamic programming model was developed to optimize multi-period investment decision in order to ultimately evaluate the results of the different investment patterns (Kennedy, 1986; Asseldonk et al., 1999). A multi-stage programming model was proposed for the optimization of investments in process integration and industrial energy technologies (Elin Svensson, 2011). Nevertheless, very few decision-making models of unconventional oil exploitation exist. Many experts and scholars predicted that the peak oil is expected to occur during the early decades of the 21st century (Hubbert, 1956; Bakhtiari, 2004;Mohr and Evans, 2008; Ugo Bardi, 2009). Conventional oil cannot continue to meet ongoing demand (Nick, 2010). Yan Luguang and other Chinese Academy of Sciences members (2006) explicitly pointed out that unconventional oil resources offer a most feasible and significant potential to supplement conventional oil resources. Michael Dale (2011) also regarded unconventional oil as the best prospects for the future. The strategic significance, production prediction, and economic and cost analysis of developing unconventional oil were primarily discussed in many papers (Mehmet Erturk, 2011; Mamdouh G. Salameh, 2003; Aurélie Méjean, 2008; David L. Greene et al., 2006; Benjamin, 2007; Mohr and Evans, 2010; Robert and Laura, 2008). Carlos de Castro (2009) stated the possible substitution of conventional with unconventional oil by using system dynamics models in his paper.

In this paper, unconventional oil is limited to extra-heavy oil, oil sands, and oil shale which are considered as the most prospective resources to substitute for conventional oil. The aim of the paper is to develop a multi-stage decision model to allocate the limited resources to maximize the total return from investment horizon in the presence of exploitation technologies. The rest of the paper is organized as follows. Section 2 analyzes the properties of oil sands, extra-heavy oil, and oil shale and provides the general formulation that will appear in the paper. Section 3 sets up a dynamic programming model of optimal investment policy over time and

discusses algorithm applied to solve the problem. Section 4 presents a numerical example which tests for the efficiency and effectiveness of the model in section 3. Section 5 presents conclusions and areas for further study.

2 Problem Statement and Complexity Analysis

Unconventional oil is rich reserves but the reserves of them are homogeneous: Canadian tar sands solely has 81% of all explored oil sands, Venezuelan heavy oil is the biggest reservoir, and US has the majority of the shale oil in the world.

2.1 Properties of Oil Sands, Extra-Heavy Oil, and Oil Shale

Canadian tar sands, as typical oil sands, consist of friable rocks of which over 76% is composed of sand, silt, clay, and natural bitumen. Bitumen is composed of carbon rich, hydrogen poor long-chain molecules with an API of less than 10^{o} and a high viscosity. So it is absolutely essential for conventional refineries to spend much more money in converting bitumen into synthetic crude oil (SCO) to be refined. Therefore, the explored oil sands reservoirs were often abandoned and undeveloped in the past (Mehmet Erturk, 2011).

Extra-heavy oil is defined as a hydrocarbon with a low API gravity that is less than 10 and a viscosity that is less than 10000cP (WEC, 2007). Extra-heavy oil is always buried more than 600m deep and the recovery rate is less than 10 percent. It's a huge expensive project that significant financial and natural gas investments are required. For example, an oil company will invest about $1 billon in the infrastructure about oilfields and pipelines not including crude oil upgrading and port infrastructure and facilities for the 0.2 Mb/d (Million barrels per day) of oil (Chen Junwu, 2009). Moreover, the environmental impacts incurred by the extra-heavy oil exploitation are devastating as well as oil sands.

Oil shale is a huge source compared to oil sands and extra-heavy oil with estimated 3 trillion barrels recoverable oil reserves globally and has been a real alternative to oil in USA (Dyni JR, 2003). There are three factors restricting the development of shale oil. Firstly, higher infrastructure investment and production costs will be afforded in shale oil production. Secondly, cost overruns always existed because of much more uncertainties. Thirdly, exploiting shale oil will incur many serious environmental problems such as poisonous gas emission, water pollution and consumption and so on. For example, U.S. Department of the Interior (2005) had concluded that there would be no free water available for shale oil exploitation and conversion if upper Colorado basin was exploited at 6 Maf/y (million acre-feet per year) by 2050.

Nonetheless, the negative and complex properties discussed above cannot stop unconventional oil exploitation moving forward because of a limited capacity of oil production and a soaring demand of oil especially in developing countries such as China and Indian. The growth of unconventional oil in the past decades has been 4–5% annually, and most of this growth depended on Canadian oil sands and Venezuelan heavy oil (Carlos de Castro et al., 2009).

2.2 Complexity of Unconventional Oil Exploitation

Each unconventional oil resource, even the same resource at different areas has its own properties. The development of unconventional oil resources requires great inputs to create a manufacturing process that integrates exploration, production, conversion, refining and other processes (Bengt Söderbergh et al., 2007). The lowest unit exploitation cost for any unconventional oil resource is much higher than the lowest production cost for conventional oil. Oil companies will increasingly look forward to unconventional oil resources to boost oil production even if unconventional oil exploitation is extremely complex, difficult and costly.

Unconventional oil exploitation is a complex system involving economic, social, technical, and environmental factors. Oil corporations always spend billions of dollars in exploring and exploiting unconventional oil as well as dealing with other important issues such as cost overrun, heavier ground and water pollution, huge quantities water consumption, higher carbon emission and much more natural gas than exploiting easy oil. For example, oil sands always contain lots of heavy metals and sulfur, and sometimes it is unavoidable to clear the huge forest and deplete vast quantities of water to exploit them. Oil companies that are prepared for oil sands exploitation have to spend much more money in open-pit mining or in situ production as well as the environmental impacts.

It is absolutely crucial for oil companies to take a rigorous method to assess the economic feasibility and profitability of unconventional oil by permission of technology and environment. Oil companies are extremely cautious to decide whether exploiting unconventional oil or not and how funds are to be apportioned among them to obtain maximum total return from the exploitation activities at the various basins suppose that we have known the amount of investible resources and proved reserves available of unconventional oil resources.

2.3 Problem Statement

An oil company has a fixed amount of Q resources and a set of M unconventional oil projects, $P_1, \ldots, P_M$, each of which depletes resources and yields return. The projects are interdependent in that an individual project's return and resource consumption is influenced by which other projects are selected (Robert L. Carraway and Schmidt, 1996). Q funds are to be apportioned among M unconventional oil projects to maximize the total profits from N years, $Y_1, \ldots, Y_N$. A new infusion of funds $I \geq 0$ within every year will be needed to continue the exploitation and operation of unconventional oil projects according to the actual situation. Maybe dividends $I \leq 0$ will be shared at the end of year when the projects operate very well. $R_k^{(j)}$ denotes the ROI at the end of year k from project j ($k=1, \ldots, N$; $j=1, \ldots, M$). The ROI of different projects in N years is shown in Table 1.

Table 1 ROI of unconventional oil exploitation in *N* years

Year / projects	Y_1	…	Y_k	…	Y_N
P_1	$R_k^{(1)}$		$R_k^{(1)}$		$R_N^{(1)}$
…					
P_i	$R_k^{(j)}$		$R_k^{(j)}$		$R_N^{(j)}$
…					
P_M	$R_M^{(j)}$		$R_k^{(M)}$		$R_N^{(M)}$

The question to be addressed here is how to allocate part of or all of the Q resources among M unconventional oil projects such that the total return from the projects are maximized simultaneously in N years. Note that the economic, social, political, and environmental issues should be taken into consideration because of the investment irreversibility of unconventional oil resources. Moreover, unconventional oil resources maybe are converted to conventional oil resources over time for the particularity of unconventional oil exploitation and the development of new exploitation techniques. The decision problem here is not a single stage investment decision to obtain maximum return at the end of the stage but a multi-stage investment portfolio to maximize total return during the whole stages. The investment decision process can be depicted in Fig.1.

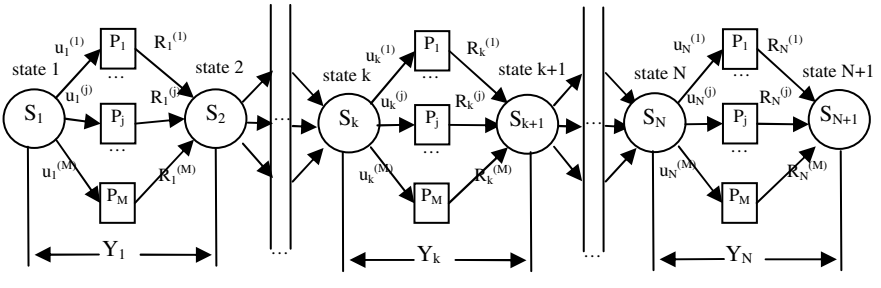


Fig 1 The state transition diagram of dynamic programming model

Here, S_k ($k=1$, …, N) presents the state at the beginning of year k. $u_k^{(j)}$ denotes a fixed amount of resources for project j at the beginning of year k ($k=1$, …, N; $j=1$, …, M).Obviously, we can see that the maximum total return over N years depends on N, the initial capital quantity Q, and decisions at the beginning of each year. The initial decision has a direct effect in subsequent years. A fundamental property of the decision process at Y_k only depends on the data up to Y_k and resources allocation at later years depend on the investment portfolio at that time. Considering the N-stage process, the total return from the N stage process will be return that the profits from the first stage plus the return from the second stage at which we have an amount $Q+(u_1^{(1)}\cdot R_1^{(1)}+\ldots+u_1^{(j)}\cdot R_1^{(j)}+\ldots u_1^{(M)}\cdot R_1^{(M)})$ of resources

allocated, continuously plus the return from stage *N*. It can be seen that the whole decision process where we repeat the above operation of allocation *N* times in succession is an *N*-stages decision problem that can be solved by the dynamic programming. Dynamic programming model can be used to help top managers in oil companies select the optimal schemes to exploit extra-heavy oil, oil sand or shale oil and obtain the maximum total return from the projects over time.

3 Mathematical Model and Algorithm

We suppose that investment capital is available, natural gas price is acceptable, and environmental issues can be dealt with well. Simultaneously, no political or environmental constraints are allowed to hinder oil production. With all these assumptions taken into consideration it is possible to select an investment policy which maximizes the total expected profits from unconventional oil projects.

Let k be stage. Y_k is defined as stage variable in this problem, $k=1, \ldots, N$. At stage k, the state variable x_k in the interval $[0, Q]$ represents the capital that can be allocated at the beginning of Y_k and the initial capital is a given amount $x_1=Q$. Let u_k be decision variable in the interval $[0, x_k]$ and $u_k^{(j)}$ denotes the resources that can be used to invest in project j at the beginning of Y_k ($k=1, \ldots, N; j=1, \ldots, M$). $I_k^{(j)}$ represents additional investment or dividends of P_j at the end of Y_k. The allowed decision set, $D_k(x_k)$, is shown as follows:

$$D_k\left(x_k\right)=\left\{u_k^{(j)} \middle| 0 \le u_k^{(j)} \le x_k\right\}\left(j=1,\ldots,M;k=1,\ldots,N\right) \tag{3.1}$$

Let x_{k+1} denote the resources available for allocation at the beginning of stage $k+1$. A state transition equation can be written as follows:

$$x_{k+1}=x_k+\sum_{j=1}^{M}\left(u_k^{(j)}\times R_k^{(j)}+I_k^{(j)}\right)\left(k=1,\ldots,N;j=1,\ldots,M\right) \tag{3.2}$$

Subject to the constraints

$$\sum_{j=1}^{M} u_k^{(j)}=x_k \tag{3.3}$$

$$x_k \ge 0$$

Note that $I_k^{(j)} \le 0$ when $I_k^{(j)}$ denotes additional investment in P_j at the end of Y_k and $I_k^{(j)} \ge 0$ when $I_k^{(j)}$ represents dividends from P_j at the end of Y_k.

We define objective function $f_k(x_k)$ which represents the maximum return obtained from an k-stage process with an investment capital x_k, for $k=1, \ldots, N$ and $x_k \ge 0$. The value function $V_k(x_k, u_k^{(j)})$ denotes the profits obtained from P_j at stage k. The maximum total return takes the following form:

$$f_k(x_k) = \max_{u_k^{(j)} \in D_k(x_k)} \left\{ \sum_{j=1}^{M} V_k(x_k, u_k^{(j)}) + f_{k+1}(x_{k+1}) \right\} \tag{3.4}$$

Where

$$V_k(x_k, u_k^{(j)}) = u_k^{(j)} \times R_k^{(j)} \tag{3.5}$$

Note that the boundary condition is

$$f_{N+1}(x_{N+1}) = 0 \tag{3.6}$$

The values V_k at earlier stage can be found by working backwards. Dynamic programming has been proven to be effective for multi-stage and single-objective or multi-objective optimization problems (Robert, 1991). The decision-making optimization problems of complex system like unconventional oil exploitation can be solved by the dynamic programming model presented in the paper.

In the mathematical optimization method of dynamic programming, backward induction is one of the main methods for solving the Bellman equation (Bellman, 1957).The algorithm developed for dynamic programming is essentially transforming the multi-stage problem into a single-stage problem. Hu Xiangpei et al. (1996) pointed out that The dynamic programming problem can be described by IGOTDS method and calculated by the improve breath-first search algorithm. Benjamin Doerr et al. (2011) presented evolutionary algorithms has been proven for solving the dynamic programming in various combinatorial optimization problems. We implement the model through backward recursion algorithm in Matlab or Lingo software to help the oil company find an optimal investment policy of unconventional oil to obtain maximum total profits from investment horizon.

4 Case Study

Assume that an oil company has founded billion tons of unconventional oil reservoirs at Ordos Basin, China. The company is going to invest $100 million in continuously exploiting extra-heavy oil, oil sands, and oil shale by permission of technology, management, human resources, and environments. The profits of the projects will be injected into the continuous exploitation and mass production of extra-heavy oil, oil sands or shale oil at the end of year. According to the experts and professionals viewpoints, the predictive ROI of extra-heavy oil, oil sands, and shale oil over time at Ordos Basin is shown in table 2.

Table 2 ROI of exploiting extra-heavy oil, oil sands, and shale oil at Ordos Basin

Year / Projects	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Extra-heavy oil	7%	15%	18%	17%	16%	15%
Oil sands	5%	8%	14%	18%	21%	22%
Shale oil	6%	5%	16%	20%	19%	18%

The decision problem is how to apportion the $100 million among extra-heavy oil, oil sands, and shale oil to maximize the total return from six years. It is a typical multi-stage decision problem that can be solved by dynamic programming. We can simplify the decision process by breaking it down into a sequence of steps and consider the problem as a *6*-stage process if let the year be stage variable.

Let the index j=1, 2, 3 denote the three potential projects, with j=1 corresponding to extra-heavy oil, j=2 corresponding to oil sands, and j=3 corresponding to shale oil. Base on the mathematical model in section 3, the *6*-stage optimization model of resources allocation problem can be modeled as:

$$f_k\left(x_k\right)=\max_{u_k^{(j)}\in D_k\left(x_k\right)}\left\{\sum_{j=1}^{3}V_k\left(x_k,u_k^{(j)}\right)+f_{k+1}\left(x_{k+1}\right)\right\} \tag{4.1}$$

s.t.

$$V_k\left(x_k,u_k^{(j)}\right)=\sum_{j=1}^{3}u_k^{(j)}\times R_k^{(j)},$$

$$x_{k+1}=x_k+\sum_{j=1}^{3}u_k^{(j)}\times R_k^{(j)},$$

$$\sum_{j=1}^{3}u_k^{(j)}=x_k,$$

$$D_k\left(x_k\right)=\left\{u_k^{(j)}\middle|0\le u_k^{(j)}\le x_k;1\le j\le 3\right\},$$

$$f_7\left(x_7\right)=0,$$

$$k=1,2,3,4,5,6,$$

$$j=1,2,3.$$

Computational result is demonstrated as follows by using backward recursion algorithm.

Table 3 The optimal solution of resources allocation in $ million

Year / Projects	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Extra-heavy oil	100	107	123.05	0	0	0
Oil sands	0	0	0	0	174.24	210.83
Shale oil	0	0	0	145.20	0	0

According to data shown in table 3, the optimal allocation of initial capital $100 million for each year is as follows:

- All of initial capital $100 million should be invested in extra-heavy oil project at the first year.
- The profits add to the capital should be invested in extra-heavy oil exploitation in year 2 and year 3.
- All of the money $145.20 million should be invested in shale oil projects in year 4.
- $174.24 should be invested in oil sands project in year 5.
- At the end of year, all of the money $210.83 million should be invested in shale oil exploitation project.

The maximum total return that the oil company will obtain from six years at Ordos Basin is $157.21 million if investment policy in table 3 is applied. Note that the annual return in case is deterministic. In reality, the ROI is changing because of fluctuating oil prices, natural gas prices, growth in proven reserves and internal cash flows, technological progress, and instable factors on regional economy and politic such that oil companies should adjust or modify investment policy timely according to the internal and external environments where they exist.

5 Conclusions

A dynamic programming model was developed in this paper to help decision makers apportion a fixed amount of resources among a set of unconventional oil projects in order to find an optimal investment policy to maximize total expected return from investment horizon. This paper provides a creative and important approach to model the decision-making process over a finite investment horizon to ensure the proper apportion of resources. The case presented the oil company that will exploit extra-heavy oil, oil sands, or shale oil at Ordos Basin can apply the investment policy shown in Table 3 to obtain maximum total return, $157.21 million, from investment horizon of six years. The results shown presented that dynamic programming model is an effective and efficient for the investment decision problems of unconventional oil exploitation. Nevertheless, many important issues such as refinery capacity, water and ground pollution, and water consumption that were not taken into consideration in the paper deserves special and serious

attention in the future rather than only the economic feasibility and maximum profits were discussed. We can apply the dynamic simulation methods taking into account oil price and technological progress over time to solve the problems for the properties of uncertainty, complexity, and investment irreversibility.

References

[1] Méjean, A.: Modelling the costs of non-conventional oil: A case study of Canadian Bitumen. Energy Policy 36, 4205–4216 (2008)
[2] Bakhtiari, A.M.S.: World oil production capacity model suggests output peak by 2006–2007. Oil and Gas Journal 102(16), 18–20 (2004)
[3] Söderbergh, B., Robelius, F., Aleklett, K.: A crash programme scenario for the Canadian oil sands industry. Energy Policy 35, 1931–1947 (2007)
[4] Sovacool, B.K.: Solving the oil independence problem: Is it possible? Energy Policy 35, 5505–5514 (2007)
[5] Doerr, B., Eremeev, A., Neumann, F., et al.: Evolutionary algorithms and dynamic programming. Theoretical Computer Science 412, 6020–6035 (2011)
[6] Bellman, R.E.: Dynamic Programming-Republished 2003. Princeton University Press, Princeton (1957)
[7] de Castro, C., Miguel, L.J., Mediavilla, M.: The role of nonconventional oil in the attenuation of peak oil. Energy Policy 37, 1825–1833 (2009)
[8] Chen, J., Li, C., Chen, X.: Oil Supplement and Substitution. Sinopec Press, Beijing (2009)
[9] Greene, D.L., Hopson, J.L., Li, J.: Have we run out of oil yet? Oil peaking analysis from an optimist's perspective. Energy Policy 34, 515–531 (2006)
[10] Dyni, J.R.: Geology and resources of some world oil-shale deposits. Oil Shale 20, 193–252 (2003)
[11] Svensson, E., Strömberg, A.-B., Patriksson, M.: A model for optimization of process integration investments under uncertainty. Energy 36, 2733–2746 (2011)
[12] Hubbert, M.K.: Nuclear energy and the fossil fuels. In: Drilling and Production Practice. American Petroleum Institute (1956)
[13] Hu, X., Qian, G., Hu, Y.: A Research on the Knowledge Representation for Discrete Dynamic Programming Model and Its IBFS Algorithm. Journal of Harbin Institute of Technology 28(3), 119–126 (1996)
[14] Kennedy, J.O.S.: Principles of dynamic optimization in resource management. Agricultural Economics 2(1), 57–72 (1988)
[15] Salameh, M.G.: Can renewable and unconventional energy sources bridge the global energy gap in the 21st century? Applied Energy 75, 33–42 (2003)
[16] van Asseldonk, M.A.P.M., Huirne, R.B.M., Dijkhuizen, A.A., Beulens, A.J.M., et al.: Dynamic programming to determine optimum investments in information technology on dairy farms. Agricultural Systems 62, 17–28 (1999)
[17] Erturk, M.: Economic analysis of unconventional liquid fuel sources. Renewable and Sustainable Energy Reviews 15, 2766–2771 (2011)
[18] Dale, M., Krumdieck, S.: Pat Bodger. Net energy yield from production of conventional oil. Energy Policy 39, 7095–7102 (2011)
[19] Mohr, S.H., Evans, G.M.: Peak oil: testing Hubbert's curve via theoretical modeling. Natural Resources Research 17(1), 1–11 (2008)

[20] Owenn, N.A., Inderwildi, O.R., King, D.A.: The status of conventional world oil reserves–Hype or cause for concern. Energy Policy 38, 4743–4749 (2010)
[21] Carraway, R.L., Schmidt, R.L.: An improved discrete dynamic programming algorithm for allocating resources among interdependent projects. Management Science 37(9), 1195–1206 (1991)
[22] Kaufmann, R.K., Shiers, L.D.: Alternatives to conventional crude oil: When, how quickly, and market driven? Ecological Economics 67, 405–411 (2008)
[23] Mohr, S.H., Evans, G.M.: Long term prediction of unconventional oil production. Energy Policy 38, 265–276 (2010)
[24] Bardi, U.: Peak oil: The four stages of a new idea. Energy 34(3), 323–326 (2009)
[25] U.S. Department of the Interior, Bureau of Reclamation Upper Colorado Region. Quality of water Colorado River Basin Progress Report No.22. U.S. Department of the Interior website (2005), http://www.usbr.gov/uc/progact/salinity/pdfs/PR22.pdf (accessed June 8, 2008)
[26] Yan, L., Chen, J., Zhou, F., et al.: Long term strategy of supplement and substitute for conventional oil in China. Advanced Technology of Electrical Engineering and Energy 25(4), 1–7 (2006)

A Proposal of Farmers Information System for Urban Markets

Tomoko Kashima, Shimpei Matsumoto, Hiroshi Iseda, and Hiroaki Ishii

Abstract. In recent years, new business models for agricultural markets have appeared. Under this perspective, we develop a new information system for urban markets to facilitate the transactions. Both sides, consumers and farmers, require certain information from markets about agricultural products. For example, consumers may make requests about the exact information of agricultural products or their safety, while farmers may want to boast with their produce. Under the considerations of such requirements at the markets, which may be conflicting, we propose a new information system to assist in the negotiation among both parties.

1 Introduction

In recent years, the number of agricultural-product markets is over the number of convenience stores in Japan. However, many farmers are experiencing a lot of problems nowadays. Farmers are rapidly aging because young people who get a agricultural job are decreasing in number. Consumers worry about the safety of the food which Farmers produce because the problem of the quantity of the agricultural

Tomoko Kashima
Kinki University, 1 Takaya Umenobe, Higashi-Hiroshima City, Hiroshima, 739-2116 Japan
e-mail: kashima@hiro.kindai.ac.jp

Shimpei Matsumoto
Hiroshima Institute of Technology, 2-1-1 Miyake, Saeki-ku, Hiroshima, 731-5193 Japan
e-mail: s.matsumoto.gk@it-hiroshima.ac.jp

Hiroshi Iseda
Institute for Agricultural Products Distribution Co., Ltd., 17-1, Owakiyama, Yamada-cho oub, Kita-ku, Kobe, 651-1101 Japan
e-mail: sptm2@kcc.zaq.ne.jp

Hiroaki Ishii
Kwansei Gakuin University, 2-1, 2-1 Gakuen, Sanda, Hyogo, 669-1337 Japan
e-mail: ishiihiroaki@kwansei.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 353–360.
springerlink.com

chemicals used for agricultural products And, there is also a growing concern that the Trans-Pacific Partnership(TPP) has a negative influence on domestic agriculture. However, these problems could be a big opportunity to try and change domestic agriculture in Japan by introducing IT. We have already developed a Menu Recommendation System which gathers data on what people like to eat, and with that data, it can make automatic recommendations for an individual[9]. This idea can also be applied in developing the Agricultural Information System for urban markets, for the benefit of farmers and the local agriculture in Japan. This system will provide information on what agricultural products are in demand by analyzing consumer consumption and market trends. With this information, the farmers can have a better idea of what crops to prioritize. This can also help stabilize the economic sustainability of farming by improving farm management. With the system at work, it will reduce oversupply and undersupply of certain agricultural products, and the stable supply-demand relationship will prevent the underpricing of agricultural products and help in stabilizing market prices.

2 Present Conditions of Japanese Gricultural

A long time ago, the Japanese agricultural system was by the self-sufficiency which makes agricultural products by itself and is eaten by itself. The agricultural products to need were created as needed. The farmhouse created agricultural products and consumers bought agricultural products directly by development of the food system. Next, agricultural products were sold from Central Wholesale Market from farmhouse to retail, and were sent to consumers. Consumers' opportunity to get to know the place-of-production information on the purchased agricultural products decreased from such change. Also at the restaurant, a customer's opportunity to get to know the detailed information of the foods used for a menu decreased. Similarly, a farmhouse's opportunity to acquire consumers' information decreased. For example, the customer needs the information what kind of products they want to buy and how season they want to get the products.

The problem of food attracted attention since about 2000. Consumers changed to become interested to food of secure and safe. Consumers began to ask for getting to know how and where agricultural products were made. Consumers came to get interested also in quantity of the agricultural chemicals of agricultural products and growing information of agricultural products. If such information about food were released to the public, a menu has added value. For example, we think that the quantity of agricultural chemicals and the information on a place of production become important. The foods of the bastard size may also be needed at a restaurant.

Many researchers did systems development and have proposed from various viewpoints[8], [5], [3].The agricultural problem is tackled also in countries other than Japan. The food safety management system in Korea is largely managed [4].The Agri-food Safety Information System planned to be upgraded continuously by connecting practical safety management information and by constructing Emergency Warning System and Crisis Coping System which can be used to promptly

cope with situation when a food accident occurs. The agricultural extension system in Tanzania has faced many problems [1], [6]. Moreover, farmers has a problem of reduction in income. Farmers are pressed for the need for change from the old selling method. The Farmers need to plan and grow the agricultural products which consumers regard as wanting rather than need to create without a plan. The Farmers need to investigate best-selling agricultural products in advance, and need to sell them in a store. However, there is a big problem here. The problem is the physical distance between consumers and farmers. Now, many of sales information of Farmers is not managed. The Farmers have not acquired a customer's sales trend.

Jensen expressed whether improvements in information impact market performance [2]. When information is limited or costly, agents are unable to engage in optimal arbitrage in fishery and agriculture. Between 1997 and 2001, mobile phone service was introduced throughout Kerala, a state in India with a large fishing industry. Using microlevel survey data, they showed that the adoption of mobile phones by fishermen and wholesalers was associated with a dramatic reduction in price dispersion, the complete elimination of waste, and near-perfect adherence to the Law of One Price.We think farmers release information and sharing information leads to the stability of the prices of agricultural products from this case.

In addition, IT-related company has begun entry in agriculture [7]. They are applied for agriculture using a clud-computing. They have focused on introducing the latest technologies (sensors, wireless networks and Cloud computing), radically revising approaches to agriculture and conducting business feasibility studies to make a hypothesis model of Cloud services that truly contribute to agriculture.

However, it is very difficult for farmers to build the large-scale system which has a network all over the country. Therefore, it is difficult for farmers to share information mutually. Moreover, it is very difficult for local farmers to do the same measure as a major company. In this research, we consider those problems, and we aims at the configuration of the farmers information system which can be used also in district agriculture (see Figure 1).

3 Problem and Solution

We think there are three problems about agricultural management. The first problem is stabilizing agricultural management. If the quantity of production is unstable, market prices would also be unstable. If supply and distribution is indefinite, the income of a farmer is also indefinite, thereby making it difficult to recruit the younger generation to engage in farming. This particular problem means that the farming industry workforce is not infused with new laborers while the existing farmers become older. In 2010, the percentage of farmers aged 65 and over was 62percent. The average age of farmers is 66 years old. The income that farmers make today is only half of what a farmer makes 20 years ago. The area of agricultural land which is not being used and cultivated is continuously increasing. The farming experience and wisdom of the expert farmers are not being put to good use and because there is no one to inherit their knowledge of

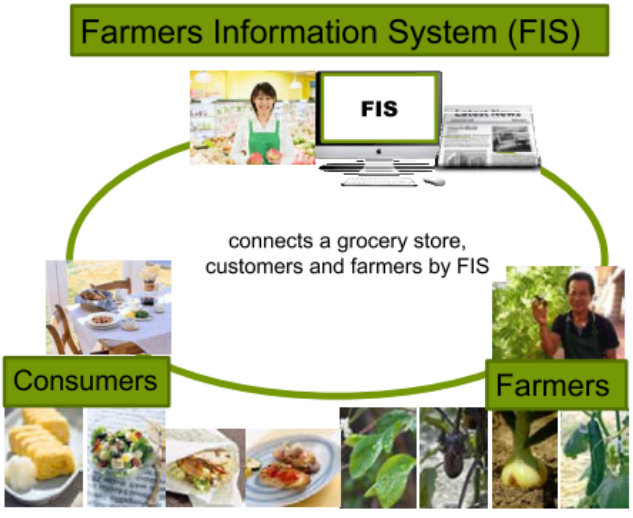


Fig. 1 Purpose of FIS

farming. If the farmers do not adjust their farming schedule to fit the demand of consumers, they automatically loose that sales opportunity. It is then very important that the farmers take into consideration what the consumers really want, to avoid not being able to sell their products or worse, sell it for a very low price. Farmers should focus on producing products that has actual consumer demand. They have to adjust their farming schedule in a way that by the time they harvest, their product is actually sought upon by consumers. For example, the farmers should only plant watermelons just in time for the summer, when it is popular.

The next problem is the unstable supply of agricultural products. Long term storage of certain agricultural products just to maintain a steady supply is difficult to do because some agricultural products have short shelf life. This problem makes it impossible for farmers to supply market demands for certain periods of time. Sometimes, certain environmental phenomenon, like the weather, affects the planting, harvesting and even delivering goods to the market difficult. Our goal is to design a system that will provide farmers information on when to plant so that the harvesting can be synchronized with actual market demands. To accurately approximate the time between the planting preparations up to the time of harvest, information on rate of seedling growth before planting, growing information and shipping time approximations until the product reaches the markets, will be used make a planting prediction. This information is then disclosed to the public via the web. This way, ordinary consumers, the food service industry, the farmers and other sectors than can benefit from this information may have access. The information on which

agricultural products the Farmers can supply and be made available to the market at a given time, may be known buy prospective buyers and consumers alike. If this is so; consumers and buyers can plan their purchases in advance. The food service industry can also take advantage of this information by being able to plan ahead the kind of food they can serve on the menu during that certain period. This system will give both farmers, consumers and the food industry sufficient time to prepare and make necessary adjustments and preparations.

The last problem is that worth of agricultural products is cheap, even though Farmers grew the agricultural products carefully. It is necessary for the farmers to provide the consumers the information related to the production of the products. Consumers are interested with the process because they need to be secure that the products they are buying are of good quality and that they were grown carefully. This agricultural information system would include pictures of the actual farmers and their farms, information on growing practices and methods will also be included, along with the information on what type of chemicals or fertilizers were used in growing the products, or what type of preservatives were added to the products, if ever there was any. The objective of the system is to provide consumers information that cannot be known by simply looking at the products. This information can be considered as a value added to the products, thereby resulting to better acceptance of the products and better market prices.

4 Farmers Information System

This system is a system which connects farmers, a customer, and a sales store (see Figure 1). The system has some pages (see Figure 2). Farmers disseminate the information on agricultural products. The system can perform an information input simply. For example, a farmer only sends the photograph taken by the cellular phone by e-mail. The system can exhibit what kind of agricultural products farmers are making. The system can exhibit the date of scheduled harvest of agricultural products. Moreover, farmers can display relief on web by displaying the photograph of products. The photograph and comment which contributed are immediately displayed on a system. Consumers can peruse the page of farmers. Consumers can also write a comment to the photograph and comment which farmers posted. They can peruse the cooking method of the purchased vegetables. Moreover, they can post the dish of boast built using the purchased vegetables. A grocery store has a duty which gives the information of farmers to consumers. Many information gathers for a system. The information is printed by paper like a newspaper (see Figure 3). Consumers can see the information on a system with a personal computer or a cellular phone. However, it is also effective to distribute information to the consumers who came to the store in paper. The new information in a system is published by paper.

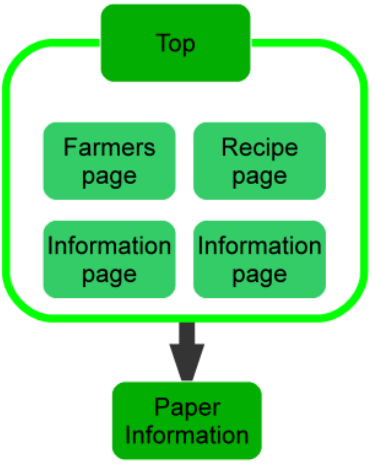


Fig. 2 Outline of Farmers Information System

Fig. 3 Paper made from the system

Fig. 4 Top of system

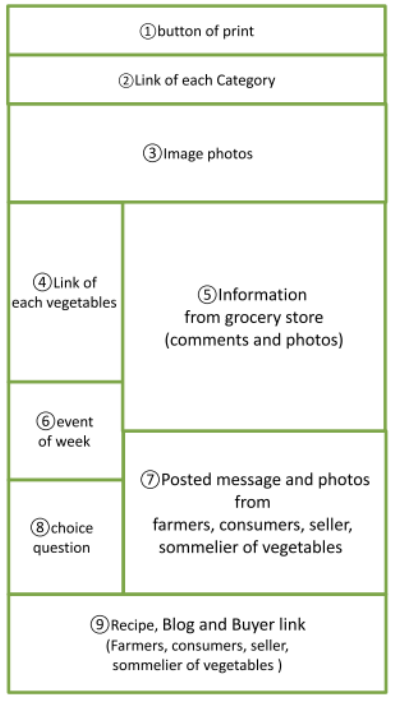


Fig. 5 Composition of top page

Figure4 and Figure5 shows the details of a system. The function of No. 1 summarizes the new information on a system to one sheet. The information on one sheet becomes like a newspaper. The function of No. 5 posts the information on the Farmers and agricultural products which the store obtained. The shop assistant can post easily from a cellular phone or a personal computer. New information is displayed on this space.The 7th function displays the information posted like the 5th. This space is posted by Farmers, consumers, seller, sommelier of vegetables, etc. The 5th is information effective in consumers and the master chief of a restaurant. Consumers can know that agricultural products are raised safely. The master chief of a restaurant can know the information on the agricultural products harvested next, when creating a menu. The 7th can acquire various information. Usual, there was no opportunity for Farmers to get to know how the agricultural products which he raised are eaten by consumers. Farmers become an opportunity to get to know consumers' comment over agricultural products. The information is useful for next cultivation for Farmers. A system has a page of Farmers, a page of information sending, a page of a recipe, an event page, etc. other than a top page. The detailed information of the agricultural products into which each agricultural products raise the page of Farmers is displayed. The page has the harvest time and the amount of agricultural chemicals of agricultural products. The calender with harvest information can be known visually.

5 Conclusion

In this paper, farmers information system for carrying out a remedy to three problems was proposed. The problem about the present agriculture was shown in this paper.Solution was proposed to those problems and the support system was developed. The structure which connects Farmers, consumers, and a store was built. The solution of the problems takes much time. However, this system will be useful for a problem solving.

As future works, the authors will support social activities towards the applications for marketing and policy planning by understanding customer's attributes and activities from stored data through the operation in the real worksite. This study hopes to construct a business support model for restaurants considering some external factors from the viewpoint of entire optimization, but the local improvement only for mall. To achieve this objective, a simulation model is needed to be constructed. The authors are planning to make a computational model based on the game theory which enables to examine the changes of overall benefit depending on collaborative relationships between restaurants.

References

1. Mattee, A.Z.: Reforming Tanzania's Agricultural Extension System: The Challenges Ahead. African Study Monographs 15, 177–188 (1994)
2. Robert, J.: The Digital Provide: Information (Technology), Market Performance and Welfare in There South Indian Fisheries Sector. The Quarterly Journal of Economics 122(3), 879–879 (2007)

3. Miyasaka, J.: Modelling and Optimization for Agricultural Production Systems. The Institute of Systems, Control and Information Engineers 54(4), 138–143 (2010)
4. Lee, K.O., Bae, Y., Kei, N.: Construction and Management Status of Agri-Food Safety Information System of Korea. Journal of the Faculty of Agriculture, Kyushu University 55, 341–348 (2010)
5. Hashimoto, M., Utsumi, T., Yukimatsu, K.: Networking System for Large Rice Field Using Independent Power Supply System and Long Distance Wireless LAN. Agricultural Information Research 16(1), 9–21 (2007)
6. Masawe, J.L.: Farming System and Agricultural Production among Small Farmers in the Uluguru Mountain Area, Morogoro Region, Tanzania. The Center for African Area Studies, Kyoto University 13, 171–183 (1992)
7. Hori, M., Kawashima, E., Yamazaki, T.: Clud Computing Applied for Agriculture and Other Fields. FUJITSU 61(3), 314–320 (2010)
8. Hoshi, T., Shiozawa, E., Shinma, K., Takaichi, M., Firufuji, M.: Development of an Application Program for Field Servers to Acquire and Leverago Production History Information in Protected Horticulture. Agricultural Information Research 16(1), 1–8 (2007)
9. Kashima, T., Matsumoto, S., Ishii, H.: Decision Support System for Menu Recommendation Using Rough Sets. International Journal of Innovative Computing, Information and Control 7(5), 2799–2808 (2011)

An Approach to the Real-Time Risk Analysis for Hazardous Material Transportation

Linlin Liu, Lijun Sun, Jianming Li, Chao Li, and Xiangpei Hu

Abstract. Real-time risk analysis is the critical point of monitoring, warning and re-routing HazardousMaterial Transportation (HMT). In this paper, we propose a real-time risk analysis approach based on Internet of Things (IoT), which dynamically evaluates the risk at discrete time pointsduring the whole transportation process.Probabilities of accidents are determined by both statistic accident data and real-time monitoring data collected by IoT techniques. To determine the potential consequences of accidents, Areal Location of Hazardous Atmospheres (ALOHA) is used to simulate the accident scenarios and estimate the influence areas. The population exposed can be calculated with the interface of Intelligent Traffic System (ITS). Furthermore, our approach provides a more accurate way to dynamically evaluate the risk of HMT.

1 Introduction

Hazardous materials, the properties of which are toxic, corrosive, explosive and flammable, often do harm to human beings, facilities and the environment. Particularly, the transportation of them is a risky activity, as explosions, side-rolling crashes and collisions usually happen during the transportation processes. Once this kind of accidents happens, it would cause deaths and environment pollutions, so the damage is inestimable.

According to the statistics of the U.S. Department of Transportation, there were 13,682 accidents relevant to hazardous materials transportation (HMT) happened in 2011, 85.5% of which occurred on the road [1]. China is one of the most productive nations of hazardous materials. To the end of 2010, there have been 22,000 enterprises that produce hazardous materials, and about 28,600 companies run relative businesses in China [2]. As the total transport amount rising year by year, the situation of supervising HMT becomes more serious.

Linlin Liu · Lijun Sun · Jianming Li · Chao Li · Xiangpei Hu
Institute of Systems Engineering, Dalian University of Technology, Dalian, China
e-mail: gloriousage@gmail.com

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 361–366.
springerlink.com

Real-time risk analysis is the basis of monitoring and warning HMT, which can be used to plan a less risky route and reduce the potential harm of accidents [3]. The Internet of Things (IoT) is an innovating technology integrating with RFID, Wireless Sensor Networks (WSN), GPS, etc.It has the ability of real-time sensing, data transporting and processing [4]. IoT has been widely used in intelligent transportation, smart grid, agriculture and supply chains, as it provides a real-time, precise and intelligent management tool [5-8]. Nowadays, the supporting techniques of IoT have becoming matured, and the infrastructure of it has been consummated day by day, which provides a new opportunity for real-time risk analysis of HMT.

In this paper, we proposed an IoT-based approach to the real-time risk analysis forHMT.The remaining part of the paper is organized as follows. In the 2nd section, the state of the art of the risk analysis in the hazardous material area is reviewed. In section 3, the real-time risk analysis framework is illustrated. Finally, we conclude this paper in section 4.

2 Literature Review

Hazardous materials can be transported by road, water and air. Due to their distinct characteristics, the methods of risk analysis under different transport modes differ a lot. Erkut etal. [9] reviewed existing risk analysis models.They thought that a comprehensively accepted model does not exist, and meanwhile the least risky routes derived from different risk analysis models may not be the same, or even vary a lot.

Bi Jun and Wang Huadong [10] proposed a risk evaluation method based on the concept of “basic routes” and used it in route planning of hazardous wastes transportation in Shenyang. Guo Xiaolin and Li Jun [11] introduced the accident gradation into the risk measurement model to improve the original model. Fabiano etal. [12] considered both the individual and societal risks, and presented a risk assessment model that is sensitive to route features and population exposed. Wu Zongzhi etal. [13] developed a quantitative risk-assessing model, and introduced the modifying factors of the possibility of hazardous accidents.

Because there are various factors that influence the transportation risk of hazardous materials, it’s very difficult to calculate the transportation risk. Although many researchers have focused on this issue, available methods still have some defects. Jiang Xuepeng etal. [14] surveyed the development of researches on risk assessment of HMT both at home and abroad. They assumed that accident rate is consistent, and transportation process is continuous, which are not rational. Fabiano etal. believed that the existing methods just focused on the inherent factors such as hazardous materials, vehicles, drivers, and neglected the weather, traffic flow, conditions and other external factors [15].

In recent years, researchers introduced sensors, GPS, GIS and other technologies to improve the analysis of risk in HMT. Verter took polygon to represent the population exposed, took polyline to represent the road links, and proposed a GIS-based risk analysis method[16]. Wei Hang considered the varied population density around the HMT route, and proposed a population-estimation-based risk analysis model[17]. Zhang Jianghua used information-diffusing theory to evaluate

the dangerous chemicals transportation accident rate, and analyzed the result of accident with GIS simulation technology[18]. Tomasoni used ALOHA to simulate the dispersion of hazardous materials, and proposed a consequence-based risk evaluation method[19].

Besides the above mentioned researches, some researchers use qualitative methods to study the risk analysis problem of HMT. Arslan applied the SWOT analysis approach to formulate the safe carriage of bulk liquid chemicals in maritime tankers[20]. Milazzo considered the potential terrorist attacks during HMT, and proposed a dynamic scenario-analysis-based risk management approach[21].

By analyzing these researches, we find that current researches have the following two drawbacks when dealing with monitoring and warning HMT:

(1) Current researches mostly evaluate the risk of the whole route before transportation.Nevertheless, the risk may vary a lot when travelling at different time points or in different parts of the route.

(2) The calculation of accident rate and result are often based on the static and historical data, which can not reveal the current transportation scenario and real-time context changes.

The proposed IoT-based approach to the real-time risk analysis for HMT in this research will overcome these drawbacks.

3 The Model for Real-Time Risk Analysis

Although there are many different kinds of risk analysis models, most of them are based on the simple idea that risk can be evaluated by the possibility of accidents together with their consequences. In this paper, we also use the classical form of risk analysis, which takes the summation of potential accident scenarios' expected consequences as the total risk for one route. To do this, we should decide two elements of these scenarios, their possibility as well as their consequences.

We define $R_t$as the location road of HMT truck at time t, then the transportation risk could be evaluated by equation (1).

$$\mathrm{D_t} = \sum_{S} f_t N_{s,t} P_s \tag{1}$$

Where D_t is the transportation risk at time t.f_t is a time-varied possibility of accidents which depends on the dynamic transportation scenarios.S is the accident scenario, which includes flaming accident, explosion, side-rolling crash, etc. $N_{s,t}$ is the population and facilities that would be affected by accident scenario S at time t, and P_s is the accident rate of scenario S. $N_{s,t} \cdot P_s$ represents the consequence of accident scenario S.

3.1 The Possibility of Potential Accidents

As mentioned above, the possibility of accident is not consistent, which changes as the transportation scenarios change. According to the previous researches, there are many factors that can affect the risk of HMT. We first divide the factors into four subgroups,

Meteorological information, including fine, rain, fogs and et al.
Road characteristics, including the lanes number, bridge or not and et al.
Traffic conditions, including high density, medium density and low density.
Status of hazard materials, including pressure, temperature, humidity and et al.

The above factors can be obtained by different IoT techniques in real time. For example, the status of hazard materials is monitored by wireless sensors, and the traffic conditions can be provided by ITS interface.

With the real-time data of these factors, we can calculate f_t as indicated in formula (2).

$$f_t = \gamma_r n_t \tag{2}$$

where n_t is the vehicle amount on road R_t at time t, γ_r is related to the accident rate of each vehicle that on R_t, and it could be calculated by equation (3).

$$\gamma_r = \gamma_0 \sum_{i=0}^{|H|} \left(|H_i| \cdot h_i \right) \tag{3}$$

Here, γ_0 stands for the basic traffic accident rate of each $year^{-1} \cdot vehicle^{-1} \cdot kilometer^{-1}$. H is set to represent the data of the subgroups mentioned above, |H| is the base of set H, and H_i is the i^{th} aspect of transportation scenario sets. Take the traffic condition set $H_{traffic}$ for example, it includes the low vehicle density, the medium vehicle density, the high vehicle density, etc. h_i is the accident rate of transportation scenario i.

3.2 The Consequence of Accident Scenarios

The accident consequence is evaluated by the population, vehicles and facilities exposed to a specific accident scenario at time t. The evaluation of consequence is based on accident scenarios. We use Areal Location of Hazardous Atmospheres (ALOHA) to simulate the chemical dispersion scenarios. Figure 1 illustrates the framework of the simulation process by ALOHA.

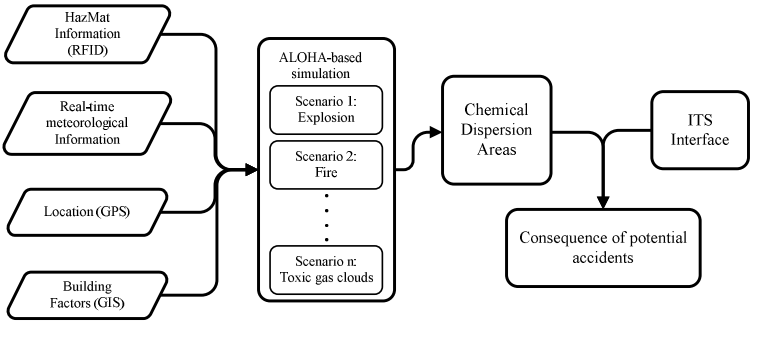


Fig. 1 The framework of the simulation process by ALOHA for accidents' consequence

Real-time input data of ALOHA is collected by techniques of IoT. The HazMat information is read from the RFID tag on the truck. The location of the truck is gotten from Global Positioning System (GPS). Building factors which will impact the dispersion of HazMat are gotten from GIS. Including the meteorological data such as wind speed and direction, ALOHA is used to estimate the HazMat dispersion areas that caused by different potential accident scenarios. By querying the carriages in the dispersion area from ITS, we can estimate the population on road.

Although wecould not get the exact population exposed in the carriages, buildings and road at specific time point using current technologies, and thus the evaluation of potential accident consequence would not accurate,however, with the development of techniques of IoT, we can almost exactly know the population exposed and thusnearly accurately evaluate the potential accident consequence.

4 Conclusions

In this paper, we proposed an IoT-based approach to the real-time risk analysis for HMT, in which the time becomes an important input variable. This approach concerns the risk at discrete time points rather than that in the whole transportation process. Techniques including RFID, WSN,and ITS are used to collect real-time data including chemical status, meteorological data, traffic conditions, etc. The probability of potential accidents is evaluated by the dynamic transportation scenarios, which can be known from the IoT monitoring data. And we used ALHOA and GIS to evaluate the consequence of a set of accident scenarios under specific transportation conditions.

Real-time risk analysis for HMT contributes to reducing the potential accident consequence and finding a transportation route with the minimized risk. Using the real-time data collected by IoT, our approach provides a more accurate way to evaluate the risk of HMT. However, IoT nowadays can't provide the exact population on roads or in buildings. Therefore, further study could focus on improving the accuracy with the techniques of IoT.

Acknowledgments This work has been partially supported by the Natural Science Foundation of China (No. 70890080, 70890083) and the Fundamental Research Funds for the Central Universities.

References

1. Incident Statistics, `http://phmsa.dot.gov/hazmat/library/data-stats/incidents` (accessed January 12, 2012)
2. State Administration of Work Safety. The twelfth five-year-plan outline for Safety Production of Hazard Materials (2011), `http://goo.gl/AZhCA` (accessed January 12, 2012)

3. Giglio, D., Minciardi, R., Pizzorni, D., et al.: Towards a decision support system for real time risk assessment of hazardous material transport on road. In: Proceeding IEMSS 2004 (International Environmental and Monitoring Software Society), pp. 1–6 (2004)
4. International Telecommunication Union. ITU Internet Reports 2005: The Internet of Things (2005) (accessed January 12, 2012)
5. Li, L.: Application of the Internet of Thing in Green Agricultural Products Supply Chain Management. In: 2011 International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 1022–1025 (2011)
6. Karnouskos, S.: The cooperative Internet of Things enabled Smart Grid. In: Proceedings of the 14th IEEE International Symposium on Consumer Electronics, ISCE 2010 (2010)
7. Dada, A., Thiesse, F.: Sensor applications in the supply chain: the example of quality-based issuing of perishables. Proceedings of Internet of Things 4952, 140–154 (2008)
8. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: A survey. Computer Networks 54, 2787–2805 (2010)
9. Erkut, E., Verter, V.: Modeling of transport risk for hazardous materials. Operations Research 46, 625–642 (1998)
10. Bi, J., Wang, H.D.: Study on the environment risk of the hazardous wastes' transportation. China Environmental Science 15, 241–246 (1995)
11. Guo, X.L., Li, J.: Transport Risk Models for Hazardous Materials Based on Accident Gradation. China Safety Science Journal 16, 35–40 (2006)
12. Fabiano, B., Currò, F., Palazzi, E., et al.: A framework for risk assessment and decision-making strategies in dangerous good transportation. Journal of Hazardous Materials 93, 1–15 (2002)
13. Wu, Z.Z., Duo, Y.Q., Liu, M., et al.: A Study on the Method of Risk Assessment of Hazardous Materials During Road transportation. Journal of Basic Science and Engineering 12, 36–44 (2004)
14. Jiang, X.P., Xu, Z.H., Deng, Y.Y., et al.: Development in Risk Study on Hazardous Chemical Materials Road Transportation. Journal of Catastrophology 21, 94–98 (2006)
15. Fabiano, B., Currò, F., Reverberi, A.P., et al.: Dangerous good transportation by road from risk analysis to emergency planning. Journal of Loss Prevention in the Process Industries 18, 403–413 (2005)
16. Verter, V., Kara, B.Y.: A GIS-based framework for hazardous materials transport risk assessment. Risk Analysis 21, 1109–1120 (2001)
17. Wei, H., Li, J.: Analysis of Population Risk in Hazardous Material Transportation as Time Varies. China Safety Science Journal 14, 95–99 (2004)
18. Zhang, J.H., Zhao, L.J.: Risk Analysis of Dangerous Chemicals Transportation. Systems Engineering-Theory & Practice 12, 117–122 (2007)
19. Tomasoni, A.M., Garbolino, E., Rovatti, M., et al.: Risk evaluation of real-time accident scenarios in the transport of hazardous material on road. Management of Environmental Quality: An International Journal 21, 695–711 (2010)
20. Arslan, O., Er, I.D.: SWOT analysis for safer carriage of bulk liquid chemicals in tankers. Journal of Hazardous Materials 154, 901–913 (2008)
21. Milazzo, M.F., Ancion, G., Lisi, R., et al.: Risk management of terrorist attacks in the transport of hazardous materials using dynamic geoevents. Journal of Loss Prevention in the Process Industries 22, 625–633 (2009)

An Intelligent Agent Framework to Manage Web Services Security

Bala M. Balachandran, Dharmendra Sharma, and Chris Peiris

Abstract. This paper describes an Intelligent Agent Framework for addressing security within a Web service environment. We have presented and discussed some of the enterprise web service security requirements It defines a comprehensive Web service security model based on intelligent software agents.

Keywords: web services, security, software agents, SOAP.

1 Introduction

Completive global market requirements are forcing organizations to adapt agile technological solutions that can dynamically respond to environmental changes. These technical solutions are based on heterogeneous software platforms. However, they are expected to communicate through interoperability features seamlessly. Web services provide a service-oriented approach to system specification, enable the componentization, wrapping and reuse of traditional applications, thereby allowing them to participate as an integrated component to an e-Business activity (Bosworth, 2001). Web services acts as a key ingredient of the Universal Business Integration Platform (UBIP) (2003). Figure 1 illustrates the components of UBIP.

UBIP consists of multiple elements. Enterprise Application Integration (EAI) is the traditional benchmark in the integration space. EAI is primarily used to integrate multiple applications on heterogeneous platforms. This view has been expanded to include both Service-Oriented Computing (SOC) and Business Process Management Systems (BPMS). In Figure 1, SOC is represented by Service Oriented Architectures (SOA), Web services, and Service Oriented Development of Applications (SODA). BPMS is depicted as Business Process

Bala M. Balachandran · Dharmendra Sharma · Chris Peiris
Faculty of Information Sciences and Engineering University of Canberra
ACT 2601, Australia
e-mail:{bala.balachandran,dharmendra.sharma}@canberra.edu.au
Chris_Peiris@hotmail.com

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 367–373.
springerlink.com

Orchestration (BPO), Business Activity Monitoring (BAM) and Business Intelligence (BI). Al though the availability of a full-fledged UBIP is still distant, active research in SOC and BPMS has established enough technological underpinnings to allow corporations to begin to explore dynamic business processes. Web services acts as major pillar to achieve this dynamic business process between multiple organizations. Web Services take advantage of the best of component-based development. Component-based object models like Distributed Component Object Model (DCOM), Remote Method Invocation (RMI), and CORBA's Internet Inter-ORB Protocol (IIOP) have been around for some time. Unfortunately, they depend on an object model–specific protocol. Web Services extend these models by communicating with Simple Object Access Protocol (SOAP) and XML to eradicate the barrier posed by the object model–specific protocols. Web services represent a new class of interoperable, web-enabled software service that utilizes open standards such as SOAP, WSDL, UDDI to bridge the gap between heterogeneous platforms.

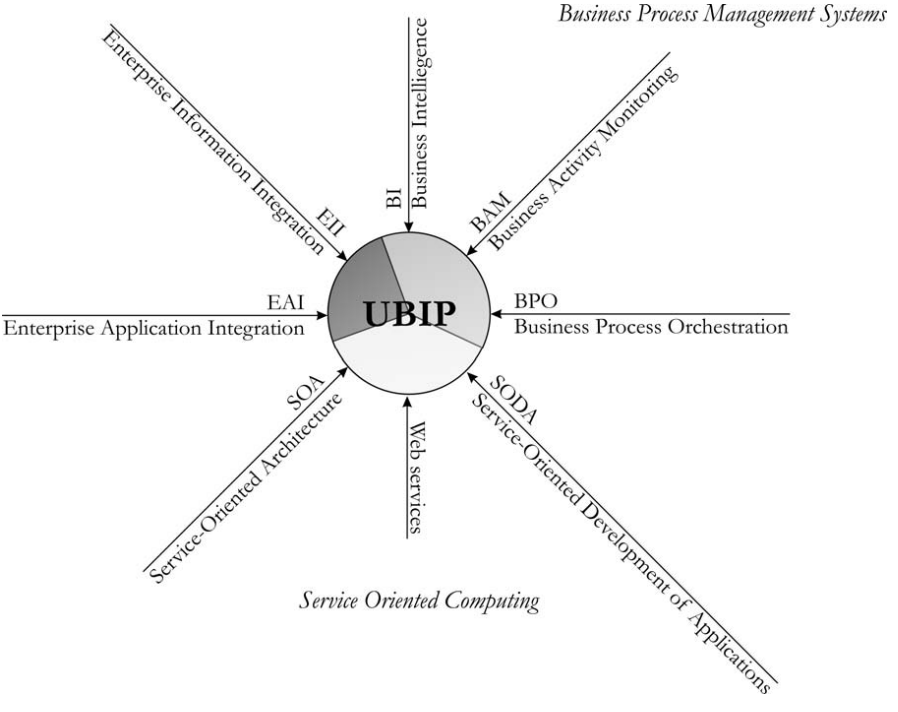


Fig. 1 Features of Universal Business Integration Platform

2 Role of Security in Enterprise Architecture

Security is a major component of any Enterprise Architecture (EA). There are several leading EA models that are used to model enterprise's information

technology capabilities. They are Zachman Enterprise Architecture Model, Federal Enterprise Architecture Framework (FEAF) and The Open Group Architecture Framework (TOGAF). In a series of articles, Malveau (2004) describes how enterprise architecture bridges the gap between business and technical strategies. He makes the point that "crucial high-level business decisions should not be constrained by technical concerns. One of the major constraints of any enterprise architecture is the security capabilities of the proposed system. We should endeavour not to restrict the business capabilities with security technical limitations. All these EA models detail security as a major component of the Enterprise Architecture.

The Zachman Framework first published in 1987, defines a logical structure for the classification and development of specific models (i.e. design artifacts) required in the management and operations of enterprises, including the development of supporting enterprise information systems. Together the inter-related models in the framework separate enterprise specific concerns (e.g. function, data, network etc.), providing a holistic enterprise. The framework is represented as a matrix where the columns of the matrix categorize knowledge significant to the enterprise into the basic interrogatives including the: what (Data); how (Function); where (Network); who (People); when (Time); and why (Motivation) of the enterprise. Security architecture is detailed under "People" under this model. Figure 2 illustrates where security architecture belongs in the wider Enterprise Architecture research scope.

The Zachman Framework	DATA *What*	FUNCTION *How*	NETWORK *Where*	PEOPLE *Who*	TIME *When*	MOTIVATION *Why*
SCOPE (Contextual) *Planner*	List of Things Important to the Business	List of Processes the Business Performs	List of Locations in Which the Business Operates	List of Organizations Important to the Business	List of Events Significant to the Business	List of Business Goals/Strategies
ENTERPRISE MODEL (Conceptual) *Owner*	Semantic Model	Business Process Model	Business Logistics System	Work Flow Model	Master Schedule	Business Plan
SYSTEM MODEL (Logical) *Designer*	Logical Data Model	Application Architecture	Distributed System Architecture	Human Interface Architecture	Processing Structure	Business Rule Model
TECHNOLOGICAL MODEL (Physical) *Builder*	Physical Data Model	System Design	Technology Architecture	Presentation Architecture	Control Structure	Rule Design
DETAILED REPRESENTATIONS (Out-of-Context) *Sub-Contractor*	Data Definition	Program	Network Architecture	Security Architecture	Timing Definition	Rule Specification
FUNCTIONING ENTERPRISE	Actual Business Data	Actual Application Code	Actual Physical Networks	Actual Business Organization	Acutal Business Schedule	Actual Business Strategy

Fig. 2 Security Architecture represented in Zachman Model

Web Service Interoperability (WS-I) has released many WS * specification to address these enterprise web service security requirements. WS-Security specification targets the security issues regarding web services on heterogeneous platforms. This Web services security model is compatible with the existing security models for authentication, data integrity and data confidentiality in common use today. As a consequence, it is possible to integrate Web services-based solutions with other existing security models? Some of the leading security models are,

- Transport Security - Existing technologies such as secure sockets (SSL/TLS) can provide simple point-to-point integrity and confidentiality for a message
- PKI - At a high level, the PKI model involves certificate authorities issuing certificates with public asymmetric keys and authorities which assert properties other than key ownership (for example, attribute authorities). Owners of such certificates may use the associated keys to express a variety of claims, including identity. The Web services security model supports security token services issuing security tokens using public asymmetric keys. PKI is used here in the broadest sense and does not assume any particular hierarchy or model.
- Kerberos - The Kerberos model relies on communication with the Key Distribution Centre (KDC) to broker trust between parties by issuing symmetric keys encrypted for both parties and "introducing" them to one another. The Web services model, again, builds upon the core model with security token services brokering trust by issuing security tokens with encrypted symmetric keys and encrypted testaments.

3 Proposed Architecture

Managing web services and its security requirements are high priority tasks for IT system designers [BIJO, 2003]. The current Web Services security architecture is driven by operating system vendors on heterogeneous platforms. The focus is on a “unified message architecture” that compliments interoperability technical limitations of heterogeneous platforms. WS-Security is the most comprehensive and widespread use of the “unified message architecture” utilizing SOAP.

We create proxy objects at the client end to direct the “message structure” from the service to the client. The proxy objects **rely on the operating system** to provide runtime assistance with security requirements (i.e. renew a security token when it has expired.) However, can we utilize intelligent agents to manage security context on heterogeneous platforms? Can we extend the capabilities of the proxy objects to assist intelligent agents to respond to environmental changes? Will it be possible to insert a new intelligent agent to derive settings from an established agent to adapt to the environment swiftly? These questions can be answers by developing an intelligent agent framework to manage web service security.

Intelligent agents are defined by Odell as autonomous, interactive (with each other), reactive to the environmental changes, goal driven and mobile (Odell, 2002) Bradshaw (1987) also refers to agents as "objects with attitude" in the sense of being objects with some additional behaviour added, for instance, mobility, inferences, etc. Objects are used as generic modelling of abstraction mechanism, independently of whether agents are implemented as objects (using object oriented programming techniques). In our research we will explore the capabilities of these intelligent agents to manage web service security. The following Figure 3 illustrates the high level functionality of the proposed framework.

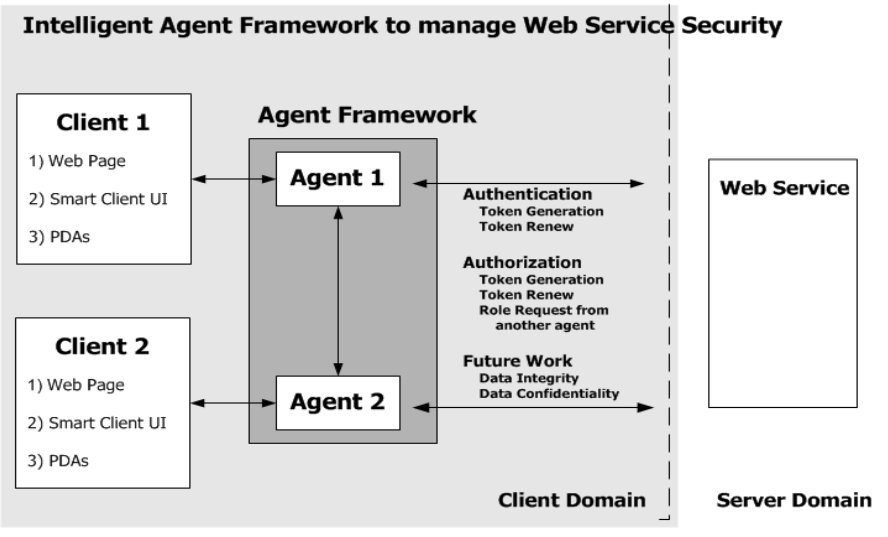


Fig. 3 Intelligent Agent Framework to Mange Web Service Security

The framework aims to utilize a set of intelligent agents to manage the authentication and authorization features of web service calls. The agent framework will reside on the clients of the web service (i.e. not at the server side with the web service). These clients can be living in heterogeneous operating systems. The agents will be primarily responsible for token management issues (i.e. token generation, token renew). We can also extend the framework to follow a "hierarchical structure of agents" to adapt to the new environment (i.e. If identical agents are deployed in the same environment, can the agents query an existing established agent to derive the security context?). This is illustrated in Figure 3 as the "double edged arrow" communication between agent 1 and agent 2.

4 Conclusions and Future Work

Web services provide a service-oriented approach to system specification, enable the componentization, wrapping and reuse of traditional applications, thereby allowing them to participate as an integrated component to an e-Business activity. We have argued that managing web services and its security requirements are high priority tasks for IT system architects and designers. Web service security requires a set of flexible, interoperable security primitives that, through policy and configuration, enable a variety of secure solutions.

To address these challenges, we have proposed an evolutionary approach to creating secure, interoperable Web services based on a set of security abstractions that unify formerly dissimilar technologies. We have proposed an intelligent agent framework for web services security.

Intelligent Agent Framework to manage Web Services Security consists of many work items and we will not be able to model all its complete functionality. Therefore, we envisage the following items as our future work.

Manage WS Federation call authentication

Manage WS Federation call authorization

Manage encryption and messaging signing by Agents (Data Integrity and Data Confidentiality)

References

Ambrosio, J.: Is now the time to manage Web services? (2004), http://www.adtmag.com/article.asp?id=9618

Bosworth, A.: Developing Web Services. In: Proceedings of the 17th International Conference on Data Engineering (2001)

Bradshaw, J.: Software Agents. MIT Press, Cambridge (1997)

Business Integration Journal Online, Universal Business Integration: An Idea Whose Time Has Come (2003), http://www.bijonline.com/PDF/matz%20march.pdf

Buhler, P., Vidal, J.M.: Towards adaptive workflow enactment using multiagent systems. Information Technology and Management Journal: Special Issue on Universal Enterprise Integration 6(1), 61–87 (2005)

IBM and Microsoft. Security in Web Services world : A Proposed Architecture and Roadmap (2002), ftp://www6.software.ibm.com/software/developer/library/ws-secmap.pdf

Malveau, R.: Bridging the Gap: Business and Software Architecture, Part 1–4 (2004), http://www.cutter.com/research/2004/edge040113.html

Odell, J.: Objects and Agents Compared. Journal of Object Technology 1(1), 41–53 (2002)

OASIS, Web Services Security: SOAP Message Security 1.1 (WS-Security 2004) OASIS Standard Specification (February 2006), http://www.oasis-open.org/committees/download.php/16790/wss-v1.1-spec-os-SOAPMessageSecurity.pdf

Peiris, C.: Creating .Net Web service (2001), http://www.15seconds.com/Issue/010430.htm
W3C, SOAP Version 1.2 (2003), http://www.w3.org/TR/soap/
Wooldridge, M.: Introduction to Multiagent Systems. John Wiley and Sons, UK (2002)
Zachman, J.A.: A Framework for Information Systems Architecture. IBM Systems Journal 26(3) (1987)

Application of Decision Rules, Generated on the Basis of Local Knowledge Bases, in the Process of Global Decision-Making

Małgorzata Przybyła-Kasperek and Alicja Wakulicz-Deja

Abstract. The paper includes issues related to process of global decision-making on the basis of knowledge which is stored in several local knowledge bases. The local knowledge bases, that are used, contain information on the same subject, but are defined on different sets of conditional attributes that are not necessarily disjoint. The paper contains a description of a multi-agent decision-making system with a hierarchical structure. In the paper a method of elimination inconsistencies in the knowledge operating on the basis of decision rules generated based on local knowledge bases was proposed.

1 Introduction

The essence of the classification problem is to make decision for a new object, using knowledge stored in the knowledge base and methods of inference. Currently, because of the large accumulation of knowledge more and more often knowledge is stored in the form of several local knowledge bases. In this situation the use of only one local knowledge base in process of decision making is not sufficient. The problem, that is considered in this paper, concerns the complex knowledge bases, in which the possibility of cooperation of local bases for making a common decision - a global decision, is very important.

The paper discusses the decision-making system that allows to make a global decision on the basis of distributed knowledge. In this system the sets of conditional attributes of different knowledge bases are not necessarily disjoint, and therefore inconsistencies of knowledge stored in different knowledge bases may occur. We understand inconsistency of knowledge to be situations in which, on the basis of

Małgorzata Przybyła-Kasperek · Alicja Wakulicz-Deja
University of Silesia, Institute of Computer Science, 41-200 Sosnowiec,
ul. Bedzińka 39, Poland
e-mail: malgorzata.przybyla-kasperek@us.edu.pl
alicja.wakulicz-deja@us.edu.pl

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 375–388.
springerlink.com

two different knowledge bases that have common attributes and for the same values for common attributes using logical implications, conflicting decisions are made. In the decision-making system, which is discussed in this paper, the process of elimination inconsistencies in the knowledge is realized in two stages. In the first stage the knowledge base that have common conditional attributes are combined into groups. In the second stage we use one of the methods of elimination inconsistencies in the knowledge in created groups of knowledge bases. Because of such structure of the system and inference being conducted in groups of knowledge bases there is a problem of conflict analysis. By conflict we mean a situation in which for the specified set of conditions on the basis of knowledge stored in different groups of knowledge bases conflicting decisions are taken. In this paper two methods of conflict analysis, that allow inference despite the presence of conflicts, were used.

In earlier work [25, 26, 27, 28, 29, 30, 31, 32, 33, 34], methods of elimination inconsistencies in the knowledge were considered, which were based on direct analysis of the knowledge stored in local knowledge bases. This paper proposes a new method of elimination inconsistencies in the knowledge, in which decision rules, generated on the basis of local knowledge bases, are used. The purpose of this paper is to compare the efficiency of inference obtained using a method of elimination inconsistencies in the knowledge that operates on the basis of decision rules with the efficiency of inference obtained using other existing methods of elimination inconsistencies in the knowledge.

The concept of taking a global decision on the basis of local decisions is also used in issues concerning the multiple model approach [7, 11, 12]. Examples of the application of this approach can be found in the literature [2, 23]. This paper describes a different approach to the global decision-making process. We assume that the set of local knowledge bases containing information from one domain is prespecified. Which means that the number of knowledge bases, as well as their sets of conditional attributes and the universe are defined before start the process of global decision-making. The paper [13], written by Ryszard Michalski and Janusz Wojtusiak, presents different approaches to information fusion. The authors consider a situation in which training examples are split to several subsets, and only aggregated values of attributes are available for each subset. In the paper [21], written by Andrzej Skowron, Hui Wang, Arkadiusz Wojna and Jan Bazan, the method of construction of a series of models under gradually relaxing conditions, which form a hierarchical structure was proposed. This method was developed to improve the accuracy of classification made on the basis of the knowledge stored in one knowledge base. The book [24], written by Vicenç Torra and Yasuo Narukawa, provides a broad overview of aggregation operators that are applied in the synthesis of judgements and information fusion. In the papers[3, 4, 9, 10] it can be found different approach to the problem of classification on the basis of several different decision tables. In methods of Distributed Data Mining (DDM), it is assumed that the data are collected and stored in different decision tables representing either horizontally or vertically partitioned. The approach considered in this paper is more general than in the DDM, because it is assumed that both the sets of attributes as well as the sets of objects can be different in different decision tables. Conflict analysis has

its roots in game theory [22] where one of the basic methods used to measure the power of agents is the Shapley-Shubik power index and the Banzhaf power index. A number of attempts have been made to solve the conflict. One of the solutions are models that describe, in a simple way, the complex structure of the conflict, these include: Zdzisław Pawlak model [14, 15] and Andrzej Skowron and Rafał Deja model [5, 19].

2 Structure of Multi-Agent System

In this paper we assume that each local knowledge base is managed by one agent. In this paper we use the definition of agent introduced by Professor Zdzisław Pawlak in [15]. We use two types of agents. The first is a *resource agent*. The resource agent has access to its local knowledge base on the basis of which he can establish the value of a local decision through the process of inference. Let Ag be a finite set of agents $Ag = \{ag_1, \ldots, ag_n\}$. For any agent $ag \in Ag$, called as the resource agent, has access to resources represented by a decision table $D_{ag} := (U_{ag}, A_{ag}, d_{ag})$, where U_{ag} is the universe and elements U_{ag} are called the objects, A_{ag} is a set of conditional attributes, d_{ag} is referred to as a decision attribute. Resource agents taking decisions on the basis of common conditional attributes form a group of agents. *A cluster of agents in* Ag is any minimal non-empty subset Ag' of Ag such that $\forall_{ag \in Ag'} \ \exists_{ag' \in Ag'} \ A_{ag} \cap A_{ag'} \neq \emptyset$ and $\forall_{ag \in Ag \setminus Ag'} \ \forall_{ag' \in Ag'} \ A_{ag} \cap A_{ag'} = \emptyset$. For each cluster a superordinate agent is defined, which is called a *synthesis agent*, as_j, where j- number of cluster. The synthesis agent, as_j, has access to knowledge that is the result of the process of inference carried out by the resource agents that belong to its subordinate group.

Definition 1. By the multi-agent decision-making system we mean $WSD_{Ag} = \langle Ag, \{D_{ag} : ag \in Ag\}, As, \delta \rangle$, where Ag is a set of resource agents; $\{D_{ag} : ag \in Ag\}$ is a set of decision tables of resource agents; As is a set of synthesis agents, $\delta : As \to 2^{Ag}$ is an injective function that assigns each synthesis agent a cluster.

Further details on the definition of multi-agent decision-making system can be found in the paper [33].

3 The Method of Aggregation of Supporting Rules

In this paper we consider the method of global decision-making in which on the basis of each decision table of resource agent, a set of decision rules is constructed. This set of decision rules is constructed using issues of the rough sets theory, which was formulated in 1982 by Professor Zdzisław Pawlak [16, 17]. In this section, the way in which sets of decision rules are generated, will be discussed.

Let $D = (U, A, d)$ be a decision table and the set $P = \{b_1, \ldots, b_k\}$ be a subset of attribute set A. Each formula α such that $\alpha = (b_1, v_1) \wedge \ldots \wedge (b_k, v_k)$, where $v_i \in V^{b_i}, i \in \{1, \ldots, k\}$ is called a *P-formula*. We say that an *object* $x \in U$ *fulfils P-formula* α *in the decision table* D, $x \models_D \alpha$, if and only if $b_i(x) = v_i$ for each $i \in \{1, \ldots, k\}$.

The expression of the form $\alpha \rightarrow (d,v)$ is called a *decision rule*, where α is a formula and $v \in V^d$. The decision rules are defined in such a way, to describe certain relationships that occur in decision table. On the basis of decision rule, it can be conducted the classification of a new object to decision class, specified by the consequent of decision rule. This classification usually proceeds as follows: if the new object meets the antecedent of the rule, the formula α, then the value of the decision attribute d to this object is equal to the decision, which occurs in the consequent of the rule. Unfortunately, such reasoning may lead from true premises to a false conclusion. That is why in many papers [1, 18, 20] different methods of generating a set of decision rules and different methods of selection rules by which we classify a new object to decision class are considered.

In this paper, in order to generate decision rules on the basis of decision tables of resource agents, the program Rough Set Exploration System (RSES [36]) was used, which was developed at Warsaw University under the direction of Professor Andrzej Skowron. This program allows to analyze data stored in a table, with using the rough set theory. The RSES program was also used to carry out initial experiments which were conducted to select a method of generation of decision rules that guarantee the best efficiency of inference of the decision system, which were considered in the paper. All available in the RSES method of generation of decision rules were considered. It was found that for the purposes considered in this paper, a method to generate a set of decision rules for resource agents, which provides the best efficiency of inference is a method that consists of two stages: generate a set of reducts for decision attribute and object; generate by local method a set of decision rules, on the basis of the set of reducts.

Now we give a definition of the reducts for decision attribute and object. Let $D = (U,A,d)$ be a decision table, $x,y \in U$ be given objects and $B \subseteq A$ be a subset of conditional attributes. We say that objects x and y are discernible by B if and only if there exists $a \in B$ such that $a(x) \neq a(y)$. A set of attributes $B \subseteq A$ is called a reduct for decision attribute and object $x \in U$ of decision table D if and only if: for any object $y \in U$ if $d(x) \neq d(y)$ and x,y are discernible by A, then they are also discernible by B; any subset of B does not satisfy the above condition.

Thus, in the first stage of method of generation decision rules the reducts are defined. In the next stage of the method for each reduct a subtable of decision table of resource agent is defined. This subtable contains conditional attributes belonging to the reduct, the universe and the decision attribute are the same as in original decision table. For such a subtable rules with the minimum number of descriptors are calculated. Further description of the methods of determining rules with the minimum number of descriptors can be found in [18, 20]. At the end, all the obtained sets of rules for each reduct of decision table of resource agent are summed to a single set of decision rules. When for each decision table the set of rules are generated then the method of elimination inconsistencies in the knowledge is used. In the method of aggregation of supporting rules, each synthesis agent has access to a set of aggregate decision rules, which are constructed on the basis of the sets of rules of resource agents, that belong to its subordinate group. Aggregated rules are created as follows. For each maximal due to inclusion set of uncontradictory decision rules of

different resource agents from one cluster, we define a new rule. By a set of decision rules of different resource agents, we understand that this set does not contain two different decision rules of one resource agent. We assume that rules $\alpha \rightarrow (d, v_1)$ and $\beta \rightarrow (d, v_2)$ are uncontradictory if and only if: $v_1 = v_2$; if in formulas α and β there exist atomic formulas with the same attribute, then the attribute values given in the atomic formulas are equal, ie $\left(\alpha \equiv (a, v_1^a) \wedge \alpha' \text{ and } \beta \equiv (a, v_2^a) \wedge \beta'\right) \Longrightarrow \left(v_1^a = v_2^a\right)$, where α', β' are formulas. The new rule is defined as follows: the rule predecessor is a conjunction of predecessors of considered decision rules of resource agents, while the rule consequent is the same as the consequent of each of the considered decision rules of resource agents. Classification, of the new object, is based on rules, selected from the set of rules of synthesis agents, for which predecessor is fulfilled by the new object. Usually, the test object is recognized by many of the rules belonging to a set of rules of one synthesis agent, and these rules classify the object into different decision classes. In this situation, each of the decision rules recognizing the test object gives vote to the decision, which occurs in the consequent of the rule. In this way, for each synthesis agent $j \in \{1, \ldots, card\{As\}\}$ an c-dimensional vector $[\mu_{j,1}(x), \ldots, \mu_{j,c}(x)]$ is calculated, where the value $\mu_{j,i}(x)$ denotes the number of votes, which the i-th decision class received when classifying the test object x, where c is the number of all decision classes. Of course, the method of analysis conflicts between rules that recognize the test object which is described above, is one possible approach to solving this problem. During initial experiments it was examined that this method ensures the highest efficiency of inference in the situation considered in the paper.

4 Method of Elimination Inconsistencies in the Knowledge and Conflict Analysis

As mentioned earlier, the main aim of this paper is to compare the efficiency of inference obtained using the method of aggregation of supporting rules, with the efficiency of inference obtained using the methods of elimination inconsistencies in the knowledge proposed in the earlier work. In the earlier work for the purpose of elimination inconsistencies in the knowledge within one cluster we used one of four methods: the method of Boolean analysis of inconsistencies, the method of aggregation of decision tables, the approximated method of aggregation of decision tables, the method of aggregation of decision tables preceded by a random editing and Hart condensing method. These methods were discussed in detail in the papers [25, 26, 27, 28, 29, 30, 31, 32, 33, 34]. The method of Boolean analysis of inconsistencies essentially consists in verifying for any two resource agents from one cluster and each decision class: whether there are relevant objects from a given decision class, which have equal values on common conditional attributes. We say that there is no inconsistency in knowledge relating to a given decision in one cluster, if for each pair of agents from the cluster above condition is satisfied. The essence of the method of aggregation of decision tables is to create a set of new decision tables on the basis of the decision tables of resource agents. The number of tables created

is equal to the number of clusters. In this method, each synthesis agent has access to resources given in the form of a decision table, which objects are constructed by combining of objects from decision tables of resource agents that belong to subordinate group, but only those objects are combined for which the values of the decision attribute and common conditional attributes are equal. The approximated method of the aggregation of decision tables is a modification of the method of aggregation of decision tables generated by applying heuristics to reduce the computational complexity. In this method, decision tables of synthesis agents are constructed on the basis of a small subset of objects relevant to the object, for which we want to take the global decision. The method of aggregation of decision tables preceded by a random editing and Hart condensing method is also a modification of the method of aggregation of decision tables. Also this time, this method has been proposed to reduce the computational complexity of the method of aggregation of decision tables by reducing the number of objects used to construct aggregate tables. This time we used the approach in which we selected from each decision table of resource agent a relatively small subset of objects representing the knowledge stored in the decision table. In order to select a subset of representative objects from the decision table of resource agent we use two methods: a random editing and Hart condensing method.

Conflict Analysis is implemented after completion of the process of elimination inconsistencies in the knowledge, because then the synthesis agents have access to the knowledge on the basis of which they can independently establish the value of a local decision to just one cluster. Two methods to resolve the conflict analysis will be used in this paper: the method of weighted voting, the method of a density-based algorithm. These methods allow the analysis of conflicts and enable to generate a set of global decisions. In the case of the method of a density-based algorithm (a algorithm DBSCAN [6] known from the literature was used), the generated set will contain not only the value of the decisions that have the greatest support of knowledge stored in local knowledge bases, but also those for which the support is relatively high. These methods were discussed in detail in the papers [28, 33].

5 Results of Computational Experiments

For the experiments the following data, which are in the UCI repository [35], were used: Soybean Data Set, Dermatology Data Set and Landsat Satellite data set.

In order to determine the efficiency of inference of the multi-agent decision-making system with respect to the analyzed data, each data set was divided into two disjoint subsets: a training set and a test set. The proposed multi-agent decision-making system allows to make global decisions on the basis of distributed knowledge, stored in several decision tables. To fully explore the possibilities and to evaluate the efficiency of inference of the multi-agent decision-making system, it is necessary to provide the knowledge stored in the form of a set of decision tables. Thus each of the training sets was divided into a set of decision tables, the division with different number of decision tables were considered. For each of the used data sets, the multi-agent decision-making system with five different structures,

versions with 3, 5, 7, 9 and 11 resource agents, were considered. Then, on the basis of each decision table, using the RSES program, the sets of decision rules were generated as discussed in Section 3. Preliminary experiments, that aim was to investigate the effect of filtering rules, shortening rules and generalization of rules[36] on the efficiency of inference in multi-agent decision-making system, were carried out. Experiments have shown that each of the above methods of modifying the rules caused a deterioration in the efficiency of inference, so in the later stages of experiments, if there was no such necessity, these methods were not used. In the case of the multi-agent decision-making systems with 3 resource agents WSD_{Ag_1} of Soybean and Dermatology data sets and the multi-agent decision-making systems with 5 resource agents WSD_{Ag_2} of Landsat Satellite data set the algorithm implementing the method of aggregation of supporting rules, operating on the basis of full set of rules has not completed analysis of the test set within 6 hours. Thus, in order to shorten the time needed to analyze the test set, for each of these decision systems the decision rules of one resource agent, who had the largest number of rules, have been shortening[36] (with a coefficient 0.99). On the basis of preliminary experiments it was found that among the methods of modifying the rules, the rules shortening is the method for which we obtain the greatest efficiency of inference in multi-agent decision-making system. In the case of the multi-agent decision-making system with 3 resource agents WSD_{Ag_1} of Landsat Satellite data set, the RSES program within 12 hours did not generate a set of rules for resource agents. Therefore, in further experiments, this decision-making system is not considered.

The measures of determining the quality of the classification are: *estimator of classification error* in which object is considered to be properly classified if the decision class used for the object belonged to the set $\hat{d}_{WSD_{Ag}}(x)$, the set of global decisions generated by the multi-agent decision-making system WSD_{Ag} for the test object x; *estimator of classification ambiguity error* in which object is considered to be properly classified if only one, correct value of the decision was generated to this object; *the average size of the global decisions sets* generated for a test set of size n $\overline{d}_{WSD_{Ag}} := \frac{1}{n}\sum_{i=1}^{n} card\{\hat{d}_{WSD_{Ag}}(x_i)\}$.

Experiments with Soybean Data Set

Tests on the Soybean data set began with the optimization of the parameters ε and *MinPts* of the method of a density-based algorithm. In Figure 1, for all five types of multi-agent decision-making systems (versions with 3, 5, 7, 9 and 11 resource agents) that were considered for Soybean data set, graphs are presented, on which the points with coordinates$(\overline{d}_{WSD_{Ag_i}}, e), i \in \{1,2,3,4,5\}$ are marked. The marked values were obtained using the algorithm implementing the method of aggregation of supporting rules and the density-based algorithm, with parameter values $\varepsilon \in (0,0.2)$, which was increased from 0 by adding a constant value. The value of parameter *MinPts* was established on the basis of initial experiments. On the graphs there are marked points, which had the greatest improvement in the efficiency of inference. These points satisfy the following condition: on the left of that point you can see a significant decrease in the value of estimator of classification error, and on the right of the point there is a slight decrease in the value of this

estimator. Moreover, it was assumed that the first coordinates of selected points (the average size of the global decisions sets) should not be much greater than 2. For each selected point, on the graph, are given the corresponding values of the parameters ε and $MinPts$. In the second step of experiments the algorithm implementing the method of aggregation of supporting rules with the weighted voting was used. Table 1 contains the results of experiments with the method of aggregation of supporting rules. In the paper [32] the results of experiments, are presented, with the Soybean data set and all structures of multi-agent decision-making system, using the algorithms implementing four methods of elimination inconsistencies in the knowledge, which have been mentioned in Section 4. For comparison, Table 1 also shows the results presented in the paper [32]. In the table the following information is given: the name of multi-agent decision-making system (System); the algorithm's symbols (Algorithm); the algorithms are identified by a code consisting of two characters. Particular pieces of code represent: a method of elimination inconsistencies in the knowledge and method of conflicts analysis, which are implemented by the algorithm. The following designations were adopted: R - the method of aggregation of supporting rules, B - the method of Boolean analysis of inconsistencies, T - the method of aggregation of decision tables, $A(m)$ - the approximated method of aggregation of decision tables, $E(l;r)$ - the

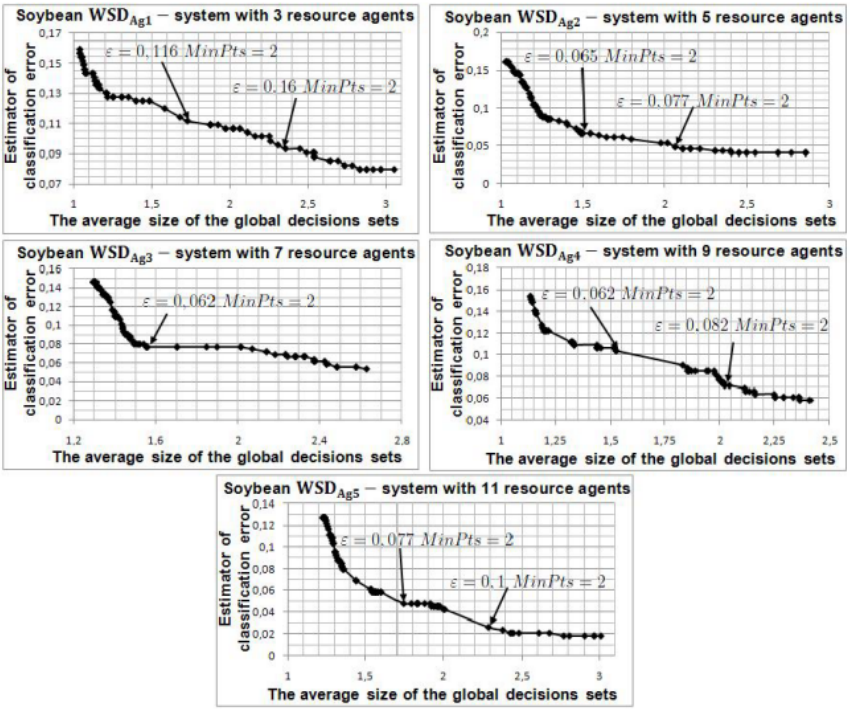


Fig. 1 Optimization of parameters of the density-based algorithm for Soybean data set

method of aggregation of decision tables preceded by a random editing and Hart condensing method, W - the method of weighted voting, $G(\varepsilon, MinPts)$ - the method of a density-based algorithm; the three measures discussed earlier $e, e_{ONE}, \overline{d}_{WSD_{Ag}}$; the time t needed to analyse a test set expressed in minutes.

Table 1 Summary of experiments results with the Soybean data set

The method of aggregation of supporting rules

System	Algorithm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag1}	$RG(0.16;2)$	0.093	0.258	0.354	22.30
	$RG(0.116;2)$	0.112	0.242	1.731	22.30
	RW	0.159	0.199	1.040	22.30
WSD_{Ag2}	$RG(0.077;2)$	0.048	0.309	2.066	2.49
	$RG(0.065;2)$	0.066	0.284	1.489	2.49
	RW	0.181	0.181	1.024	2.49
WSD_{Ag3}	$RG(0.062;2)$	0.077	0.277	1.557	8.04
	RW	0.189	0.239	1.295	8.04
WSD_{Ag4}	$RG(0.082;2)$	0.072	0.255	2.021	0.19
	$RG(0.062;2)$	0.104	0.221	1.524	0.19
	RW	0.178	0.242	1.069	0.19
WSD_{Ag5}	$RG(0.1;2)$	0.027	0.303	2.290	0.08
	$RG(0.077;2)$	0.048	0.287	1.739	0.08
	RW	0.157	0.226	1.231	0.08

The methods that do not require the generation of decision rules

System	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag1}	$TG(0.0025;2)$	0.027	0.282	1.699	7.43
	$A(1)G(0.0025;2)$	0.027	0.295	2.005	0.03
	$E(200;119)G(0.00275;2)$	0.086	0.396	2.019	0.03
	BW	0.146	0.223	1.146	0.01
	TW	0.112	0.186	1.114	7.43
WSD_{Ag2}	$TG(0.006;2)$	0.024	0.266	1.705	3.51
	$A(1)G(0.00575;2)$	0.035	0.327	1.968	0.02
	$A(1)G(0.0035;2)$	0.093	0.242	1.335	0.02
	$E(200;119)G(0.007;2)$	0.110	0.412	2.008	0.03
	BW	0.144	0.215	1.141	0.02
	TW	0.122	0.178	1.096	3.51
WSD_{Ag3}	$TG(0.0065;2)$	0.008	0.287	1.665	2.42
	$A(1)G(0.00575;2)$	0.008	0.306	1.968	0.03
	$A(1)G(0.00375;2)$	0.024	0.274	1.559	0.03
	$E(200;119)G(0.00675;2)$	0.056	0.325	1.876	0.03
	$E(200;119)G(0.0045;2)$	0.071	0.295	1.551	0.03
	BW	0.114	0.253	1.215	0.02
	TW	0.085	0.199	1.152	2.38
WSD_{Ag4}	$TG(0.012;2)$	0.029	0.324	1.843	2.37
	$A(1)G(0.0108;2)$	0.058	0.319	1.755	0.03
	$E(200;119)G(0.0124;2)$	0.118	0.391	1.973	0.03
	BW	0.090	0.194	1.176	0.02
	TW	0.082	0.181	1.122	2.37
WSD_{Ag5}	$TG(0.0148;2)$	0.032	0.316	2.008	1.54
	$TG(0.0132;2)$	0.035	0.293	1.673	1.52
	$A(1)G(0.0128;2)$	0.029	0.309	1.755	0.03
	$A(1)G(0.0084;2)$	0.090	0.253	1.298	0.03
	$E(200;119)G(0.0148;2)$	0.126	0.425	1.960	0.03
	BW	0.109	0.242	1.274	0.02
	TW	0.098	0.207	1.141	1.52

On the basis of the results of experiments given in Table 1 the following conclusions can be drawn. For all considered types of multi-agent decision-making systems algorithms $TG(\varepsilon, MinPts)$ and $A(m)G(\varepsilon, MinPts)$ obtain better results than the algorithm $RG(\varepsilon, MinPts)$. Only for multi-agent decision-making system with 11 resource agents WSD_{Ag_5} the algorithm $RG(\varepsilon, MinPts)$ obtains a lower value of estimator of classification error (e), but at the cost of increased value of average size of the global decisions $(\overline{d}_{WSD_{Ag}})$. The algorithm $RG(\varepsilon, MinPts)$ usually obtains much better results than the algorithm $E(l,r)G(\varepsilon, MinPts)$. The only exception is multi-agent decision-making system with 3 resource agents WSD_{Ag_1}. For this system the algorithm $E(l,r)G(\varepsilon, MinPts)$ obtained better results than the algorithm $RG(\varepsilon, MinPts)$. In the case of multi-agent decision-making system with 7 resource agents WSD_{Ag_3} the algorithm $RG(\varepsilon, MinPts)$ obtained results similar to results obtained for the algorithm $E(l,r)G(\varepsilon, MinPts)$. The algorithms BW and TW obtain better results than the algorithm RW.

Experiments with Dermatology Data Set

Tests on the Dermatology data set began with the optimization of the parameters ε and $MinPts$ of the method of a density-based algorithm. Optimization of the parameters was carried out in the same way as for the Soybean data set. In the second step of experiments the algorithm implementing the method of aggregation of supporting

rules with the weighted voting was used. Table 2 contains the results of experiments with the method of aggregation of supporting rules. In the paper [32] the results of experiments, are presented, with the Dermatology data set and all structures of multi-agent decision-making system, using the algorithms implementing four methods of elimination inconsistencies in the knowledge, which have been mentioned in Section 4. For comparison, Table 2 also shows the results presented in the paper [32]. Designations used in Table 2 are the same as in Table 1.

Table 2 Summary of experiments results with the Dermatology data set

The method of aggregation of supporting rules

System	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag1}	$RG(0.028;2)$	0.045	0.127	1.155	12.20
	RW	0.091	0.091	1	12.20
WSD_{Ag2}	$RG(0.117;2)$	0.009	0.264	1.564	8.40
	$RG(0.084;2)$	0.018	0.200	1.245	8.40
	RW	0.145	0.145	1	8.40
WSD_{Ag3}	$RG(0.09;2)$	0	0.118	1.136	0.10
	$RG(0.033;2)$	0.009	0.073	1.064	0.10
	RW	0.100	0.100	1	0.10
WSD_{Ag4}	$RG(0.197;2)$	0.009	0.200	1.400	0.08
	$RG(0.051;2)$	0.018	0.064	1.055	0.08
	RW	0.127	0.127	1	0.08
WSD_{Ag5}	$RG(0.143;2)$	0	0.127	1.245	0.05
	$RG(0.128;2)$	0.009	0.118	1.173	0.05
	RW	0.136	0.136	1	0.05

The methods that do not require the generation of decision rules

System	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag1}	$TG(0.002;2)$	0	0.091	1.100	11.24
	$TG(0.001;2)$	0.009	0.018	1.009	11.24
	$A(1)G(0.0015;2)$	0	0.091	1.100	0.01
	$A(1)G(0.0005;2)$	0.009	0.018	1.009	0.01
	$E(200;80)G(0.00325;2)$	0.029	0.325	1.607	0.02
	$E(200;80)G(0.0025;2)$	0.051	0.258	1.329	0.02
	BW	0.018	0.018	1	0.01
	TW	0.018	0.018	1	11.23
WSD_{Ag2}	$TG(0.00325;2)$	0	0.073	1.073	0.40
	$TG(0.00075;2)$	0.009	0.027	1.018	0.40
	$A(1)G(0.00575;2)$	0	0.273	1.427	0.01
	$A(1)G(0.0015;2)$	0.055	0.155	1.100	0.01
	$E(200;96)G(0.0095;2)$	0.014	0.294	1.501	0.02
	$E(200;96)G(0.006;2)$	0.031	0.208	1.250	0.02
	BW	0.045	0.045	1	0.01
	TW	0.027	0.027	1	0.39
WSD_{Ag3}	$TG(0.0065;2)$	0	0.118	1.136	3.42
	$TG(0.0055;2)$	0.009	0.100	1.100	3.42
	$A(1)G(0.0095;2)$	0.009	0.255	1.318	0.01
	$A(1)G(0.006;2)$	0.027	0.191	1.191	0.01
	$E(200;96)G(0.01505;2)$	0.011	0.318	1.518	0.02
	$E(200;96)G(0.0084;2)$	0.036	0.213	1.221	0.02
	BW	0.027	0.045	1.018	0.01
	TW	0.082	0.091	1.009	3.42
WSD_{Ag4}	$TG(0.00725;2)$	0	0.136	1.145	6.09
	$TG(0.00275;2)$	0.027	0.091	1.073	6.09
	$A(1)G(0.012;2)$	0.009	0.200	1.291	0.01
	$A(1)G(0.00725;2)$	0.018	0.155	1.145	0.01
	$E(200;112)G(0.0168;2)$	0.009	0.283	1.523	0.03
	$E(200;112)G(0.0104;2)$	0.029	0.199	1.239	0.03
	BW	0.036	0.045	1.009	0.01
	TW	0.055	0.064	1.009	6.06
WSD_{Ag5}	$TG(0.0145;2)$	0	0.155	1.236	1.52
	$TG(0.00625;2)$	0.009	0.100	1.100	1.52
	$A(1)G(0.00925;2)$	0	0.136	1.191	0.01
	$A(1)G(0.00625;2)$	0.009	0.118	1.118	0.01
	$E(200;112)G(0.0207;2)$	0.009	0.260	1.503	0.03
	$E(200;112)G(0.0135;2)$	0.021	0.181	1.226	0.03
	BW	0.027	0.036	1.009	0.01
	TW	0.036	0.036	1	1.52

On the basis of the results of experiments given in Table 2 the following conclusions can be drawn. In the case of multi-agent decision-making systems with 3 and 5 resource agents WSD_{Ag_1}, WSD_{Ag_2} algorithms $TG(\varepsilon, MinPts)$ and $A(m)G(\varepsilon, MinPts)$ obtain better results than the algorithm $RG(\varepsilon, MinPts)$. In the case of multi-agent decision-making system with 7 resource agents WSD_{Ag_3} the algorithm $RG(\varepsilon, MinPts)$ obtains significantly better results than the algorithm $A(m)G(\varepsilon, MinPts)$ and slightly better results than the algorithm $TG(\varepsilon, MinPts)$. In the case of multi-agent decision-making system with 9 resource agents WSD_{Ag_4} the algorithm $RG(\varepsilon, MinPts)$ obtains low value of estimator of classification error ($e = 0.018$) with very low value of average size of the global decisions ($\overline{d}_{WSD_{Ag}} = 1.055$), which is much lower than for the algorithms $TG(\varepsilon, MinPts)$ and $A(m)G(\varepsilon, MinPts)$.

In the case of multi-agent decision-making system with 11 resource agents WSD_{Ag5} the algorithm $RG(\varepsilon, MinPts)$ obtains the same value of estimator of classification error (e) as algorithms $TG(\varepsilon, MinPts)$ and $A(m)G(\varepsilon, MinPts)$ at a slightly higher value of average size of the global decisions $(\overline{d}_{WSD_{Ag}})$. For all considered types of multi-agent decision-making systems the algorithm $RG(\varepsilon, MinPts)$ obtains better results than the algorithm $E(l,r)G(\varepsilon, MinPts)$. The algorithms BW and TW obtain better results than the algorithm RW.

Experiments with Landsat Satellite Data Set

As mentioned earlier in the case of Landsat Satellite data set and multi-agent decision-making system with 3 resource agents WSD_{Ag1} the RSES program within 12 hours did not produce a set of rules for resource agents, and therefore this type of multi-agent decision-making system is not considered below.

Tests on the Landsat Satellite data set began with the optimization of the parameters ε and $MinPts$ of the method of a density-based algorithm. Optimization of the parameters was carried out in the same way as for the Soybean data set. In the second step of experiments the algorithm implementing the method of aggregation of supporting rules with the weighted voting was used. Table 3 contains the results of experiments with the method of aggregation of supporting rules. In the paper [32] the results of experiments, are presented, with the Landsat Satellite data set and all structures of multi-agent decision-making system, using the algorithms implementing four methods of elimination inconsistencies in the knowledge, which have been mentioned in Section 4. For comparison, Table 3 also shows the results presented in the paper [32]. Designations used in Table 3 are the same as in Table 1.

Table 3 Summary of experiments results with the Landsat Satellite data set

The method of aggregation of supporting rules

System	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag2}	$RG(0.55;2)$	0.059	0.262	1.743	48.00
	$RG(0.314;2)$	0.100	0.212	1.249	48.00
	RW	0.164	0.164	1	48.00
WSD_{Ag3}	$RG(0.528;2)$	0.035	0.278	1.737	19.23
	$RG(0.272;2)$	0.071	0.202	1.245	19.23
	RW	0.163	0.163	1	19.23
WSD_{Ag4}	$RG(0.576;2)$	0.036	0.279	1.739	14.05
	$RG(0.312;2)$	0.072	0.216	1.251	14.05
	RW	0.166	0.166	1	14.05
WSD_{Ag5}	$RG(0.65;2)$	0.031	0.300	1.749	3.27
	$RG(0.284;2)$	0.063	0.218	1.254	3.27
	RW	0.155	0.155	1	3.27

The methods that do not require the generation of decision rules

System	Algorytm	e	e_{ONE}	$\overline{d}_{WSD_{Ag}}$	t
WSD_{Ag2}	$A(1)G(0.0031;2)$	0.012	0.396	1.744	3.08
	$A(1)G(0.0012;2)$	0.047	0.221	1.230	3.08
	$E(200;528)G(0.0049;2)$	0.024	0.438	1.720	8.37
	$E(200;528)G(0.0021;2)$	0.073	0.262	1.250	8.37
	BW	0.096	0.096	1	3.05
WSD_{Ag3}	$A(7)G(0.0046;2)$	0.012	0.381	1.716	4.55
	$A(7)G(0.0022;2)$	0.047	0.232	1.255	4.55
	$E(200;660)G(0.0064;2)$	0.021	0.450	1.709	9.51
	$E(200;660)G(0.0027;2)$	0.071	0.267	1.255	9.51
	BW	0.099	0.099	1	3.09
WSD_{Ag4}	$A(10)G(0.0057;2)$	0.020	0.348	1.698	6.32
	$A(10)G(0.00285;2)$	0.052	0.216	1.229	6.32
	$E(200;396)G(0.0102;2)$	0.030	0.412	1.701	9.26
	$E(200;396)G(0.0045;2)$	0.077	0.265	1.256	9.26
	BW	0.093	0.093	1	3.18
WSD_{Ag5}	$A(1)G(0.00585;2)$	0.010	0.400	1.733	3.25
	$A(1)G(0.00285;2)$	0.048	0.230	1.259	3.25
	$E(200;396)G(0.0135;2)$	0.024	0.415	1.728	9.59
	$E(200;396)G(0.00555;2)$	0.070	0.263	1.258	9.59
	BW	0.096	0.096	1	3.16

On the basis of the results of experiments given in Table 3 the following conclusions can be drawn. For all considered types of multi-agent decision-making systems the algorithm $A(m)G(\varepsilon, MinPts)$ obtains better results than the algorithm $RG(\varepsilon, MinPts)$. In the case of multi-agent decision-making systems with 5 and 7

resource agents WSD_{Ag_2}, WSD_{Ag_3} the algorithm $E(l,r)G(\varepsilon, MinPts)$ obtains better results than the algorithm $RG(\varepsilon, MinPts)$. In the case of multi-agent decision-making systems with 9 and 11 resource agents WSD_{Ag_4}, WSD_{Ag_5} and the value of average size of the global decisions $(\overline{d}_{WSD_{Ag}})$ approximately equal to 1.25 the algorithm $RG(\varepsilon, MinPts)$ obtains lower values of estimator of classification error (e) than the algorithm $E(l,r)G(\varepsilon, MinPts)$. The algorithm BW obtains better results than the algorithm RW.

6 Conclusion

In conclusion we note that for almost all considered types of multi-agent decision-making systems the algorithms $TG(\varepsilon, MinPts)$ (the method of aggregation of decision tables and the density-based algorithm) and $A(m)G(\varepsilon, MinPts)$ (the approximated method of aggregation of decision tables and the density-based algorithm) obtained better results than the algorithm $RG(\varepsilon, MinPts)$ (the method of aggregation of supporting rules and the density-based algorithm). The exception is the multi-agent decision-making system with 7 resource agents WSD_{Ag_3} of Dermatology data set, where the algorithm $RG(\varepsilon, MinPts)$ obtained better results. In addition, algorithms BW (the method of Boolean analysis of inconsistencies and the weighted voting) and TW (the method of aggregation of decision tables and the weighted voting) obtained better results than the algorithm RW (the method of aggregation of supporting rules and the weighted voting) for all considered types of multi-agent decision-making systems. On the basis of these results it can be concluded that the greater efficiency of inference provides aggregation of decision tables of resource agents than aggregation of decision rules generated on the basis of decision tables of resource agents. For most considered types of multi-agent decision-making systems the algorithm $RG(\varepsilon, MinPts)$ obtained better results than the algorithm $E(l,r)G(\varepsilon, MinPts)$ (the method of aggregation of decision tables preceded by a random editing and Hart condensing method and the density-based algorithm). Exceptions are the multi-agent decision-making system with 3 resource agents WSD_{Ag_1} of Soybean data set, and multi-agent decision-making systems with 5 and 7 resource agents WSD_{Ag_2}, WSD_{Ag_3} of Satellite Landsat data set. On the basis on these results it can be concluded that the greater efficiency of inference provides aggregation of decision rules generated on the basis of decision tables of resource agents than aggregation of subsets of representative objects selected from decision tables of resource agents.

References

1. Bazan, J.: Metody wnioskowań aproksymacyjnych dla syntezy algorytmów decyzyjnych. PhD thesis, Supervisor A. Skowron (1998)
2. Bazan, J., Peters, J., Skowron, A., Nguyen, H., Szczuka, M.: Rough set approach to pattern extraction from classifiers. In: Electronic Notes in Theoretical Computer Science, vol. 82. Elsevier Science Publishers (2003)

3. Caragea, D.: Learning Classifiers from Distributed, Semantically Heterogeneous, Autonomous Data Sources. PhD thesis, Iowa State University (2004)
4. Chawla, N., Eschrich, S., Hall, L.: Creating Ensembles of Classifiers. In: IEEE International Conference on Data Mining, pp. 580–581 (2001)
5. Deja, R.: Conflict analysis, Rough Sets; New Developments. In: Polkowski, L. (ed.) STUDFUZZ. Physica-Verlag (2000)
6. Ester, M., Kriegel, H., Sander, J., Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, Portland, pp. 226–231 (1996)
7. Gatnar, E.: Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji. PWN, Warszawa (2008)
8. Hart, P.: The condensed nearest neighbor rule. IEEE Transactions on Information Theory 16, 515–516 (1968)
9. Jiang, W., Zhang, X., Cohen, A., Ras, Z.: Multiple Classifiers for Different Features in Timbre Estimation. Advances in Intelligent Information Systems, 335–356 (2010)
10. Kargupta, H., Park, B., Johnson, E., Sanseverino, E., Silvestre, L., Hershberger, D.: Collective Data Mining From Distributed Vertically Partitioned Feature Space. In: Workshop on Distributed Data Mining. International Conference on Knowledge Discovery and Data Mining (1998)
11. Koronacki, J., Ćwik, J.: Statystyczne systemy uczace sie. EXIT, Warszawa (2008)
12. Kuncheva, L.: Combining pattern classifiers methods and algorithms. John Wiley & Sons (2004)
13. Michalski, R., Wojtusiak, J.: The Distribution Approximation Approach to Learning from Aggregated Data. Reports of the Machine Learning and Inference Laboratory, MLI 08-2. George Mason University, Fairfax (2008)
14. Pawlak, Z.: An Inquiry Anatomy of Conflicts. Journal of Information Sciences 109, 65–78 (1998)
15. Pawlak, Z.: On conflicts. Int. J. of Man-Machine Studies 21, 127–134 (1984)
16. Pawlak, Z.: Rough sets. International Journal of Information & Computer Sciences 11, 341–356 (1982)
17. Pawlak, Z.: Rough Sets: Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Boston (1991)
18. Pawlak, Z., Skowron, A.: A rough set approach for decision rules generation. ICS Research Report 23/93, Warsaw University of Technology. In: Proceedings of the IJCAI 1993 Workshop W12: The Management of Unicertainly in AI, France (1993)
19. Skowron, A., Deja, R.: On Some Conflict Models and Conflict Resolutions. Romanian Journal of Information Science and Technology 3(1-2), 69–82 (2002)
20. Skowron, A.: Boolean Reasoning for Decision Rules Generation. In: Komorowski, J., Raś, Z.W. (eds.) ISMIS 1993. LNCS, vol. 689, pp. 295–305. Springer, Heidelberg (1993)
21. Skowron, A., Wang, H., Wojna, A., Bazan, J.: Multimodal Classification: Case Studies. T. Rough Sets, 224–239 (2006)
22. Straffin, P.: Game Theory and Strategy. Mathematical Association of America, New York (1993)
23. Ślezak, D., Wrólewski, J., Szczuka, M.: Neural network architecture for synthesis of the probabilistic rule based classifiers. In: Electronic Notes in Theoretical Computer Science, vol. 82. Elsevier Science Publishers (2003)
24. Torra, V., Narukawa, Y.: Modeling Decisions Information Fusion and Aggregation Operators. Springer, Berlin (2007)
25. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Hierarchiczny system wieloagentowy. ZN Pol. Śl. Studia Informatica 28, 4(74), 63–80 (2007)

26. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Hierarchical Multi-Agent System. Recent Advances in Intelligent Information Systems, pp. 615–628. Academic Publishing House EXIT (2009)
27. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Podejmowanie decyzji globalnej z zastosowaniem hierarchicznego systemu wieloagnetowego oraz algorytmu mrókowego. ZN Pol. Śl. Studia Informatica 30, 2A(83), 213–227 (2009)
28. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Global decisions Taking on the Basis of Multi-Agent System with a Hierarchical Structure and Density-Based Algorithm. Concurrency, Specification and Programming CS&P 2009, pp. 616–627. Uniwersytet Warszawski (2009)
29. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Podejmowanie decyzji globalnych z zastosowaniem wieloagentowego systemu decyzyjnego oraz rodziny klasyfikatorów. In: Wakulicz-Deja, A. (red.) Systemy Wspomagania Decyzji, pp. 83–92. Instytut Informatyki Uniwersytetu Ślaskiego, Katowice (2010)
30. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Multi-Agent Decision Taking System. Fundamenta Informaticae 101(1-2), 125–141 (2010)
31. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Wieloagentowy system decyzyjny - porównanie metod. ZN Pol. Śl. Studia Informatica 31, 2A(89), 173–188 (2010)
32. Przybyła-Kasperek, M.: Wieloagentowy system decyzyjny - analiza konfliktów. PhD thesis, Supervisor: Wakulicz-Deja, A. (2010)
33. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Application of the method of editing and condensing in the process of global decision-making. Fundamenta Informaticae 106 (1), 93–117 (2011)
34. Wakulicz-Deja, A., Przybyła-Kasperek, M.: Wieloagentowy system decyzyjny - analiza porównawcza metod. In: Wakulicz-Deja, A. (red.) Systemy Wspomagania Decyzji, pp. 73–100. Instytut Informatyki Uniwersytetu Ślaskiego (2011)
35. http://www.ics.uci.edu/mlearn/MLRepository.html
36. http://logic.mimuw.edu.pl/rses/

Application of the Cournot and Stackelberg Mixed Duopoly Models for Mexican Real Estate Market

Vitaliy V. Kalashnikov, Daniel Flores Curiel, and Javier González

Abstract. This paper investigates the Mexican real estate market, especially the economy class homes. This sector plays an important role for Mexican social stability. Authors investigate the Cournot and Stackelberg Mixed Duopoly models where there is a competition between a state-owned public firm maximizing domestic social surplus, and a private firm, which maximizes its own profit. Such models are used for modeling markets with low number of participants and high impact of government.

1 Introduction

The crediting system has a crucial role for the society in Mexico due to many reasons. Some of them are: the deficit of apartments in the renting segment. Next, the problem of poverty is still very important to Mexico, as around 50% of population have monthly incomes below 3000-4000 pesos. This part of population depends on the state-owned projects of the construction and financing the real estate.

One of the most important projects is INFONAVIT system. Every officially working citizen has a saving account, to where the part of paid tax is being transferred. On house/apartment purchase, one can use saved money and the rest of real estate's volume is financed with credit. This credit has governmental guaranties in case of a default.

The computational game theoretic modeling tool offered in this paper is composed as a mixed complementarity problem (MCP). Numerical experiments are programmed in computing language GAMS.

Vitaliy V. Kalashnikov · Daniel Flores Curiel · Javier González
Univesidad Autonoma de Nueo Leon, Facultad de Economia, FE, UANL,
Nuevo León, Mexico
e-mail: kalashnikov_de@yahoo.de

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 389–396.
springerlink.com

This paper is structured as follows: the introduction is followed by the section with mathematical description. The analysis of obtained average credit rates and conclusions will finish the paper.

At the present time, the mortgage system (both private and public sectors) is offering changing and fixed interest rates. Fixed rates are around 11-16% and special rates are about 7-8%, but being calculated on the base of minimal salaries. If inflation rate is higher than salary increments, the credit giving bank counts a loss in real money value. If minimum salary grows faster than the inflation, the bank has an increase in actual credit value. During last years, as the figure 1 illustrates, in average there is an increase of the minimal salary and inflation state quite constant (4.5%) (Villar (2007)).

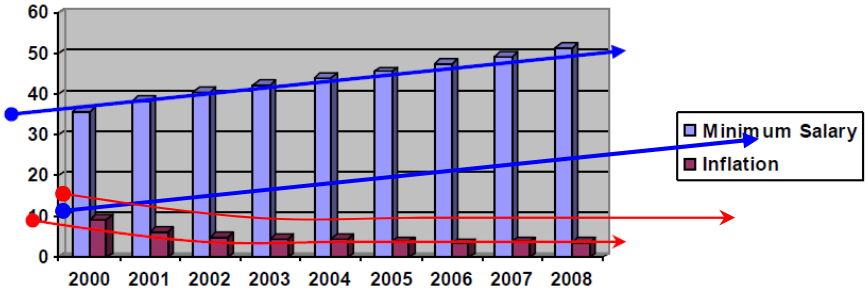


Fig. 1 The relationships between minimal salaries and inflation rate in Mexico.

The reason for using Cournot-Nash and Stackelberg Mixed models is mainly discussed in Section 2.1. These models are very popular for modeling markets with oligopolistic structures.

2 Formulation of Models

2.1 Formulation of the Model and Cournot-Nash Equilibrium

An examination of mixed oligopolies, in which social surplus-maximizing public firms compete against profit-maximizing private firms, have become increasingly popular in recent years. For pioneering works on mixed oligopolies, see Merrill and Schneider (1966), Harris and Wiens (1980), and Bös (1986, 1991). Excellent surveys can be found in Vickers and Yarrow (1988), De Fraja and Delbono (1990), Nett (1993).

The interest in mixed oligopolies is high because of their importance in economies worldwide (see Matsushima and Matsumura, 2003) for analysis of "herd behavior" by private firms in many branches of the economy in Japan.

In the paper by Matsumura (2003), the author investigates mixed duopoly and analyzes a desirable role (either leader or follower) of the public firm. This paper examines the desirable roles of both the foreign private agent and the domestic

public firm. The following figure gives a graphical interpretation for different types of models (STRA: Cournot Nash model, STACK : Stackelberg, COMP : perfect competition).

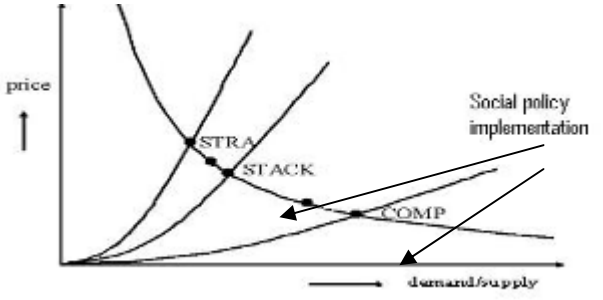


Fig. 2 Social policy: lower interest rates under same type of the model.

A mathematically extended version of applied model by Kalashnikov et al. (2007) was published in the Proceedings of the 2nd ICICI Conference.

As discussed in sections above, market with two firms offering a homogeneous product (credit) . Let G represent the total credit output, and $p(G)$ denote an inverse demand function, i.e. the interest rate consumer has to pay. Let $q_{i,t}$ ***i=1,2, t=1,…,n***, denote the output of firm ***i*** using the capital source ***t***. Same list of sources both private and public firms has been assumed. At last, let $c_{i,t}(q_{i,t})$ stand for the production cost by firm ***i*** for the technology ***t***. As G is the total output, one has

$$G = \sum_{t=1}^{n} (q_{1,t} + q_{2,t}) \quad . \tag{1}$$

Firm 1 is a foreign private firm, which maximizes its own profits, and firm 2 is a domestic public firm that maximizes domestic social surplus. Domestic social surplus S is the sum of consumer surplus and profits of firm 2, and is given by:

$$S(G, q_2) = \int_0^G p(x)dx - \sum_{t=1}^{n} \left[p(G)q_{1,t} + c_{2,t}(q_{2,t}) \right] \; ; \tag{2}$$

here, of course $q_2 = \sum_{t=1}^{n} q_{2,t}, q_1 = \sum_{t=1}^{n} q_{1,t}, q_1 = G - q_2$.

The profit of firm 1 is given by: $\Pi(G, q_1) = p(G)q_{1,t} - \sum_{t=1}^{n} c_{1,t}(q_{1,t})$. (3)

First consider the classical Cournot equilibrium, i.e., a vector $Z = (G, q_{1,1}, ..., q_{1,n}, q_{2,1}, ..., q_{2,n}) \in R_+^{2n+1}$, such that:

$$G = \sum_{t=1}^{n} \sum_{i=1}^{2} q_{i,t}\ , \tag{4}$$

$$\begin{aligned} & q_{1,t} \geq 0, \quad \phi_{1,t} \equiv c_{1,t}^{'}(q_{1,t}) - q_{1,t} p^{'}(G) - p(G) \\ & \geq 0, \ q_{1,t}\phi_{1,t} = 0 \ \forall t \end{aligned}; \tag{5}$$

$$\begin{aligned} & q_{2,t} \geq 0, \quad \phi_{2,t} \equiv c_{2,t}^{'}(q_{2,t}) + (G - q_{2,t}) p^{'}(G) \\ & - p(G) \geq 0, \quad q_{2,t}\phi_{2,t} = 0, \quad \forall t \end{aligned}. \tag{6}$$

Problem (4) – (6) is a standard complementarity problem. Therefore, applying analogous methods as in Kalashnikov et al. (2007), one can obtain the existence and the uniqueness of the equilibrium.

2.2 *Stackelberg Model with Leadership of Domestic (Public) Firm*

First, in this Section, the game where firm 2 (public one) is the leader is examined.

Firm 2 chooses its output volume $q_{2,t}$, $t = 1, ..., n$, and firm 1 (private one) chooses $q_{1,t}$, $t = 1, ..., n$ after having observed $q_{2,t}$, $t = 1, ..., n$, so as to maximize its net profit

$$\Pi(G, q_1) = p(G)q_1 - \sum_{t=1}^{n} c_{1,t}(q_{1,t})\ , \tag{7}$$

where $q_2 = \sum_{t=1}^{n} q_{2,t}, q_1 = \sum_{t=1}^{n} q_{1,t}, q_1 = G - q_2$.

Let $q_1 = q_1(q_2) \geq 0$ be the (optimal) reaction function of firm **1**; that is, the value that satisfies the equality:

$$\begin{aligned} &\phi_1(q_1(q_2)) \equiv \frac{\partial}{\partial q_1}\Pi(G, q_1(q_2)) \le 0, \text{ and} \\ &\phi_1(q_1(q_2)) \cdot q_1(q_2) = 0 \end{aligned} \tag{8}$$

Definition 2.2.1. A Stackelberg equilibrium (with the domestic firm as a leader and the foreign firm as a follower) is the vector $Z = (G^{F,L}, q_{1,1}^F(Q^L), ..., q_{1,n}^F(Q^L), Q^L) \in R_+^{n+2}$ such that

$$G^{F,L} = \sum_{t=1}^{n} \left\{ q_{1,t}^F(Q^L) \right\} + Q^L, \tag{9}$$

$$Q^L \in Arg \max \left\{ S_1(Q) \;\middle|\; Q \ge 0 \right\}, \tag{10}$$

$$q_1^F\left(Q^L\right) = \arg\max \left\{ \Pi\left(G^{F,L}, q_1\right) \;\middle|\; q_1 \ge 0 \right\}. \tag{11}$$

where $Q^L = \sum_{t=1}^{n} q_{2,t}^L$ - total output of the domestic company and $q_1^F = \sum_{t=1}^{n} q_{1,t}^F(Q^L)$ - optimal response of the private firm. The following three relationships

$$G = Q^L + q_1^F, \tag{12}$$

$$\begin{aligned} &q_{1,t}^F \ge 0, \quad \psi_{1,t} \equiv c_{1,t}^{'}(q_{1,t}^F) - q_{1,t}^F p^{'}(G) \\ &-p(G) \ge 0, \quad q_{1,t}^F \psi_{1,t} = 0 \;\forall \end{aligned}; \tag{13}$$

$$\begin{aligned} &q_{2,t}^L \ge 0, \quad \psi_{2,t} \equiv c_{2,t}^{'}(q_{2,t}^L) + (G - q_{2,t2,t}^L) p^{'}(G) \\ &-p(G) \ge 0, \quad q_{2,t}\psi_{2,t} = 0, \quad \forall t \end{aligned} \tag{14}$$

define a complementarity problem, analogous to the one discussed in Kalashnikov et al. (2007).

2.3 Stackelberg Model with Leadership of Foreign (Private) Firm

Now consider the game where firm 1 (foreign private firm) is a leader. Firm 1 chooses it $q_{1,t}$, $t = 1, ..., n$ and firm 2 (domestic or public supplier) chooses

$q_{2,t}$, $t=1,...,n$ after having observed $q_{1,t}$, $t=1,...,n$, so as to maximize domestic social surplus:

$$S(G,q_2)=\int_0^G p(x)dx-\sum_{t=1}^{n}\left[\, p(G)q_{1,t}+c_{2,t}(q_{2,t})\right] \tag{15}$$

where $q_2=\sum_{t=1}^{n}q_{2,t}, q_1=\sum_{t=1}^{n}q_{1,t}, q_1=G-q_2$.

Definition 2.3.1. A Stackelberg equilibrium (with the foreign firm as a leader and the domestic firm as a follower) is the vector $Z=(G^{L,F},Q_1^L,q_{2,1}^F(Q_1^L),...,q_{2,n}^F(Q_1^L))\in R_+^{n+2}$ such that

$$G^{L,F}=Q_1^L+q_2^F\left(Q_1^L\right), \tag{16}$$

$$Q_1^L\in Arg\max\left\{\Pi_1\left(Q_1\right)\middle|\ Q_1\geq 0\right\}, \tag{17}$$

$$q_2^F\left(Q^L\right)=\arg\max\left\{S\left(G^{L,F},q_2\right)\middle|\ q_2\geq 0\right\}. \tag{18}$$

where $Q^L=\sum_{t=1}^{n}q_{1,t}^L$ - total output of the private firm and $q_2^F=\sum_{t=1}^{n}q_{2,t}^F(Q^L)$ - optimal response of the domestic company. Following three sentences

$$G=Q^L+q_2^F, \tag{19}$$

$$\begin{aligned} &q_{1,t}^L\geq 0,\ \ \varphi_{1,t}\equiv c_{1,t}^{'}(q_{1,t}^L)-q_{1,t}^L p^{'}(G)\\ &-p(G)\geq 0,\ \ q_{1,t}^F\varphi_{1,t}=0\ \forall t \end{aligned}; \tag{20}$$

$$\begin{aligned} &q_{2,t}^F\geq 0,\ \ \varphi_{2,t}\equiv c_{2,t}^{'}(q_{2,t}^F)+(G-q_{2,t2,t}^F)p^{'}(G)\\ &-p(G)\geq 0,\ \ q_{2,t}^F\varphi_{2,t}=0,\ \ \forall t \end{aligned} \tag{21}$$

define a complementarity problem, analogous with to the one discussed in Kalashnikov et al. (2007).

2.4 Perfect Competition

Besides, it is instructive to compare the values of the produced electricity in cases of oligopolistic behaviors with the total value of the perfect competition equili-

brium, that is, when the producers ignore variations in price and solve the following complementarity problem: Find a q_1, q_2 such that

$$\begin{aligned} &\beta(q_i) \equiv c_i'(q_i) - p(G(q_i)) \geq 0, \\ &\text{and } \beta(q_i) q_i = 0, \quad \forall i = 1,2, \end{aligned} \tag{22}$$

and the total value of the electricity produced in the perfect competition scenario is equal $G = q_1 + q_2$. The inverse demand is:

$$G = D_o \cdot \left(\frac{p(G)}{p_o} \right)^{\sigma} \quad , \tag{23}$$

where D_o reference credit demand, p_o weighted average interested rate on the credit, σ - elasticity parameter.

3 Theoretical Results

In order to illustrate obtained theoretical results, numerical experiments for the classical Cournot-Nash model and the mixed Cournot-Nash one were performed. For the inverse demand function, the function of form (23) has been chosen. Elasticity parameter σ was assumed to be equal to 0.9., which refer to dependence of the credit volumes on the interest rate (Figure 3) .This figure represents the increase in time of the total credit volume (billions of pesos), the inflation rates, and the average interest rates respectively.

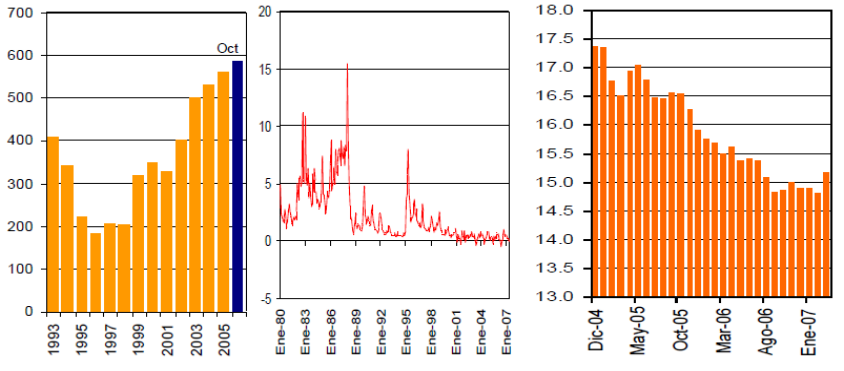


Fig. 3 Total credit volume, the inflation rate, average interest. (Source: Bank of Mexico)

The socially-oriented model offers the credit rates under the existing rates – around 6.7%. The classical model yields the equilibrium rate equal to 9.3%, which is still lower than actually offered rates. This difference can be explained by underestimation of non-payment risks, which is a goal of the further analysis.

4 Conclusion

In the above analysis, we investigated three different types of equilibrium in the duopoly with a private (foreign) agent aiming at maximization of its own profit, and a domestic firm maximizing domestic social surplus. After presenting the Cournot equilibrium in the above-described model, we examined two versions of Stackelberg game, with the private firm as a leader and domestic one as a follower, and vice versa.

In order to compare the equilibrium volumes in various scenarios we introduce the concepts of a weak and a strong firm, in dependence on the sign of the agent's optimal reaction function's derivative at the Cournot equilibrium. With such a characteristic, it turns out that if the inverse demand function is convex, then the state owned market agent is always weak, and vice versa: if the inverse demand function is concave, then the domestic agent is always strong.

For the Stackelberg equilibrium with the public owned agent as a leader, we obtain that the production volume by the leader (and hence, the total cleared market volume) is higher than that in the Cournot equilibrium, if the private firm (the follower) is strong. Otherwise, if the private agent is weak, then the total cleared market volume is lower with the domestic producer as a leader (Kalashnikov et al. (2007)).

Acknowledgements. This research was supported by the Mexico National Counsel on Science and Technology (CONACyT) within the project SEP-2011 Siencia Basica 169765; PROMEP 103.5/11/4330; PAICYT 464-10.

References

1. Bös, D.: Public Enterprise Economics. North-Holland, Amsterdam (1986)
2. Fraja, G., Delbono, F.: Game theoretic models of mixed oligopoly. Journal of Economic Surveys 4, 1–17 (1990)
3. Matsumura, T.: Stackelberg mixed duopoly with a foreign competitor. Bulletin of Economic Research 55, 275–287 (2003)
4. Merril, W., Schneider, N.: Government firms in oligopoly industries: A short-run analysis. Quarterly Journal of Economics 80, 400–412 (1966)
5. Nett, L.: Mixed oligopoly with homogeneous goods. Annals of Public and Cooperative Economics 64, 367–393 (1993)
6. Isac, G., Bulavsky, V.A., Kalashnikov, V.V.: Complementarity, Equilibrium, Efficiency and Economics, 468 p. Kluwer Academic Publishers, Dordrecht (2002)
7. Cordero, E., Kalashnikov Jr., V.V.: Stackelberg equilibrium in a mixed duopoly. In: Proceedings of the 2nd International Conference on Innovative Computing, Information and Control (ICICIC 2007), Kumamoto, Japan, September 05-07, 4 p 2007)
8. Pal, D.: Endogenous trimming in a mixed oligopoly. Economic Letters 61, 181–185 (1998)
9. Villar, A.J.: "Políticas de vivienda en México: De la Constitución de 1917 a la globalización. Universidad Politécnica de Madrid, Tesis Doctoral (Noviembre 2007)

Bayesian Network Based Prediction Algorithm of Stock Price Return

Yi Zuo, Masaaki Harada, Takao Mizuno, and Eisuke Kita

Abstract. This paper describes the stock price return prediction using Bayesian network. The Bayesian network gives the probabilistic graphical model that represents previous stock price returns and their conditional dependencies via a directed acyclic graph. When the stock price is taken as the stochastic variable, the Bayesian network gives the conditional dependency between the past and future stock prices. In the present algorithm, the stock price return distribution is transformed to the discrete values set by using Ward method, which is one of the clustering algorithms. The Bayesian network gives the conditional dependency between the past and future stock prices. The stock price is determined from the discrete value set of the stock prices so that its occurrence probability is maximized. Finally, the present algorithm is compared with the traditional time-series prediction algorithms in the TOYOTA motor corporation stock price prediction. The present algorithm show 20% better than the time-series prediction algorithms.

1 Introduction

For predicting the stock price, several time-series prediction algorithms have been studied by many researchers[1, 2, 3, 4, 5]. Autoregressive (AR) model[1, 2], Moving Average (MA) model[3], Autoregressive Moving Average (ARMA) model[4] and Autoregressive Conditional Heteroskedasticity (ARCH) model[5] are very popular time-series prediction algorithms. Recently, the extensions of ARCH model have been presented; Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model[6] and Exponential General Autoregressive Conditional Heteroskedastic (EGARCH) model[7].

Yi Zuo · Masaaki Harada · Takao Mizuno
Graduate School of Information Science, Nagoya University

Eisuke Kita
Graduate School of Information Science, Nagoya University
e-mail: kita@is.nagoya-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 397–406.
springerlink.com

The time-series forecast algorithms represent the future stock price by the linear combination of the past stock prices and the error term following to the normal distribution. Recent studies in econophysics point out that the stock price distribution does not follow the normal distribution[8]. Therefore, the error term modeled according to the normal distribution may not predict the stock price accurately. Therefore, the stock price prediction by using Bayesian Network is presented in this study. In the present algorithm, the stock price is transformed to the discrete values set by using Ward method, which is one of the clustering algorithms. The Bayesian network[9, 10] gives the probabilistic graphical model that represents previous stock price returns and their conditional dependencies via a directed acyclic graph. The future stock price can be predicted by using the network. The stock price is determined from the discrete value set of the stock prices so that its occurrence probability is maximized. Finally, the present algorithm is applied for prediction of the TOYOTA Motor Corporation stock price. The results are compared with the time-series prediction algorithm.

The remaining part of the manuscript is as follows. In section 2, time-series prediction algorithms are explained briefly. Bayesian network algorithm and the present algorithm are explained in sections 3 and 4. Numerical results are shown in section 5. The results are summarized again in section 6.

2 Prediction Algorithms

2.1 Time-Series Prediction

2.1.1 AR Model[1, 2]

The notation r_t denotes the stock price return at time t. In AR model AR(p), the return r_t is approximated with the previous return r_{t-i} $(i = 1, \cdots, p)$ and the error term u_t as follows

$$r_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + u_t \tag{1}$$

where α_i $(i = 0, \cdots, p)$ is the model parameter. The error term u_t is a random variable from the normal distribution centered at 0 with standard deviation equal to σ^2.

2.1.2 MA Model[3]

In MA model MA(q), the stock price return r_t is approximated with the previous error term u_{t-j} $(j = 1, \cdots, q)$ as follows

$$r_t = \beta_0 + \sum_{j=1}^{q} \beta_j u_{t-j} + u_t \tag{2}$$

where β_j $(j = 0, \cdots, q)$ is the model parameter.

2.1.3 ARMA Model[4]

ARMA model is the combinational model of AR and MA models. In ARMA model ARMA(p,q), the stock price return r_t is approximated as follows.

$$r_t = \sum_{i=1}^{p} \alpha_i r_{t-i} + \sum_{j=1}^{q} \beta_j u_{t-j} + u_t \tag{3}$$

2.1.4 ARCH Model[5]

In ARCH model ARCH(p,q), the stock price return r_t at time t is approximated as follows

$$r_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + u_t. \tag{4}$$

The error term u_t is given as follows

$$u_t = \sigma_t z_t \tag{5}$$

where $\sigma_t > 0$ and the function z_t is a random variable from the normal distribution centered at 0 with standard deviation equal to 1.

The volatility σ_t^2 is approximated with

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^{q} \beta_j u_{t-j}^2. \tag{6}$$

2.1.5 Determination of Model Parameters

In each model, the model parameters p and q are taken from $p = 0, 1, \cdots 10$ and $q = 0, 1, \cdots, 10$. Akaike's Information Criterion (AIC)[11] is estimated in all cases. The parameters p and q for maximum AIC are adopted.

2.2 Bayesian Network Model

In the time-series algorithms, the stock price distribution is assumed to be according to the normal distribution. Recent studies, however, point out that the distribution of the stock price fluctuation does not follow the normal distribution[8]. Figure 1 shows the stock price return of TOYOTA Motor Corporation. This figure is plotted with the P/E ratio as the horizontal axis and the data distribution as the vertical axis. The bar chart and the solid line denote the actual stock price distribution and the normal distribution. We notice from this figure that the actual distribution is very far from the normal distribution and especially, the standard deviation around $\pm\sigma$ and $\pm 3\sigma$ is conspicuous. As a result, the normal distribution may not forecast the stock price accurately. For overcoming this difficulty, Bayesian network is applied for the stock price forecast in this study.

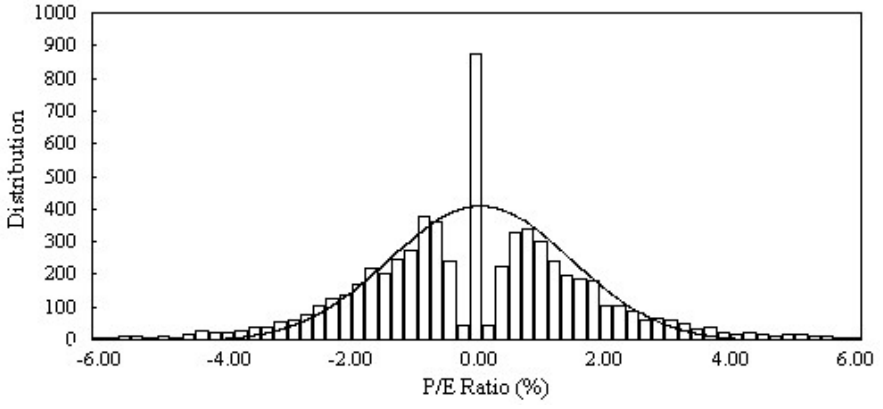


Fig. 1 P/E ratio frequency distribution of TOYOTA motor corporation stock price

The Bayesian network represents the conditional dependencies of random variables via a directed acyclic graph. The Bayesian network model can predict the stock price without the normal distribution error model. Figure 2 shows the relationship between the AR model, one of the ordinary models, and the Bayesian networkD While, in the AR model, the stock price is approximated with the linear model of the stock prices, the present method gives the nonlinear and stochastic model of the stock prices by using Bayesian network.

3 Bayesian Network

3.1 Conditional Probability Table

If the random variable x_i depends on the random variable x_j, the variable x_j and x_i are called as a parent and a child nodes, respectively. Their dependency is

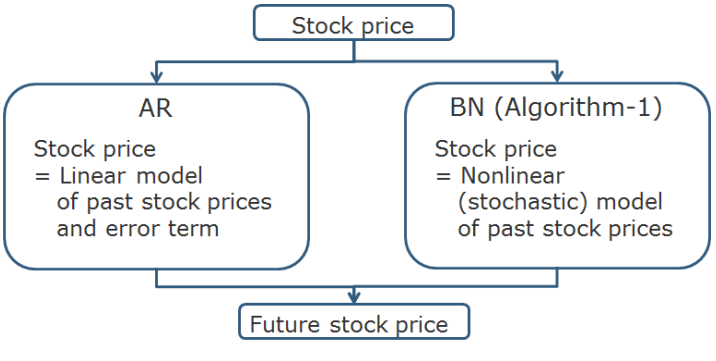


Fig. 2 Comparison of time-series and present algorithms

represented with $x_j \to x_i$. If the child node x_i has more than one parent nodes, the notation $Pa(x_i)$ denotes the parent node set for x_i. Conditional dependency probability between x_j and x_i is represented with $P(x_i|x_j)$, which means the conditional probability of x_i given x_j.

The strength of relationships between random variables is quantified with the conditional probability table (CPT). The notation Y^m and X^l denote the m-th state of $Pa(x_i)$ and the l-th state of x_i, respectively. The conditional probability table is given as follows.

$$\begin{array}{c} P(X^1|Y^1), P(X^2|Y^1), \cdots, P(X^L|Y^1) \\ \vdots \qquad\qquad\qquad \vdots \\ P(X^1|Y^M), P(X^2|Y^M), \cdots, P(X^L|Y^M) \end{array}$$

where M and L are total numbers of the states of $Pa(x_i)$ and x_i, respectively.

3.2 Determination of Graph Structure

In this study, the Bayesian networks are determined by using K2 algorithm[9, 10] with K2Metric[9, 10, 12] as the estimator of the network.

K2Metric is given as follows[12, 13]

$$K2 = \prod_{i=1}^{N} \prod_{j=1}^{M} \frac{(L-1)!}{(N_{ij}+L-1)!} \prod_{k=1}^{L} N_{ijk}! \tag{7}$$

where

$$N_{ij} = \sum_{k=1}^{L} N_{ijk}. \tag{8}$$

The notation N, L, and M denote total number of nodes, total numbers of states for x_i and $Pa(x_i)$, respectively. The notation N_{ijk} denotes the number of samples of $x_i = X^k$ when $Pa(x_i) = Y^j$.

K2 algorithm determines the network from the totally ordered set of the random variables which is summarized as follows.

1. $i = 1$
2. Set the parents set $Pa(x_i)$ for the note x_i to an empty set.
3. Estimate K2 metric S_{best} of the network composed of x_i and $Pa(x_i)$.
4. For $j = i+1, \cdots, N$,
   a. Add x_j to $Pa(x_i)$.
   b. Estimate K2 metric S of the network composed of x_i and $Pa(x_i)$.
   c. Delete x_j from $Pa(x_i)$ if $S < S_{best}$.
5. $i = i+1$
6. Go to step 2 if $i \leq N$.

3.3 Probabilistic Reasoning

When the evidence e of the random variable is given, the probability $P(x_i|e)$ is estimated by the marginalization with the conditional probability table[14].

The probability $P(x_i = X^l|e)$ is given by the marginalization algorithm as follows.

$$P(x_i = X^l|e) = \frac{\sum_{j=1,j\neq i}^{N}\sum_{x_j=X^1}^{X^L} P(x_1,\cdots,x_i = X^l,\cdots,x_N,e)}{\sum_{j=1}^{N}\sum_{x_j=X^1}^{X^L} P(x_1,\cdots,x_N,e)} \tag{9}$$

where the notation $\sum_{x_j=X^1}^{X^L}$ denotes the summation over all states $X^1, X^2, \cdots X^L$ of the random variable x_j.

4 Prediction Algorithm

4.1 Process

The process of the prediction algorithm is summarized as follows.

1. The stock price return is discretized according to the Ward method.
2. The Bayesian network B is determined by the set of the discretized stock prices.
3. The stock price return is predicted by using the network B.

4.2 Discrete Value Set of Stock Price Return

The stock price return r_t is defined as follows

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100 \tag{10}$$

where the notation P_t denotes the closing stock price at time t.

When the stock price return is transformed into the set of some clusters with the Ward method, the notation C_l and c_l denote the cluster and its center.

The discrete value set of the stock price return is given as follows

$$\{r^1, r^2, \cdots, r^L\} = \{c_1, c_2, \cdots, c_L\} \tag{11}$$

where the notation L denotes the total number of the discretized values.

4.3 Ward Method

Ward method[15] defines clusters so that the Euclid distances from samples to the cluster centers are minimized. The notation z, C_i and c_i denote the sample, the cluster and its center, respectively. The estimator is given as

$$D(C_i, C_j) = E(C_i \cup C_j) - E(C_i) - E(C_j) \tag{12}$$

Fig. 3 Total order of stock price returns

$$E(C_i) = \sum_{z \in C_i} d(z, c_i)^2 \tag{13}$$

where the notation $d(z, c_i)$ denotes the Euclid distance between z and c_i.

4.4 Stock Price Return Prediction

For determining the network B by K2 algorithm, the total order of the random variable sets is necessary. The stock price return are totally ordered according to the order of their time-series (Fig. 3).

Once the network B is determined, the stock price r_t is determined from the discrete value set of the stock prices so that its occurrence probability $P(r^l|B)$ is maximized.

$$r_t = \arg\max_{r^l}(P(r^l|B)) \tag{14}$$

Table 1 Discrete number versus AIC on forecast error of TOYOTA Motor Corp. stock return

Discrete number L	AIC
2	2.8452
3	2.5892
4	2.3913
5	2.3243
6	2.5309
7	2.7240
8	2.4935
9	2.5941
10	2.5994

5 Numerical Example

TOYOTA motor corporation stock price is considered as an example. Bayesian network B is determined from the TOYOTA motor corporation stock price from February 22nd 1985 to December 30th 2008. Then, the network is applied for predicting the stock price return from January 1st to March 30th, 2009.

5.1 Number of Clusters

We will discuss the effect of the cluster number to the prediction accuracy. The stock price return data are clustered into 2 to 10 clusters by Ward method. The AIC values of the discrete value sets are compared in Table 1. Table shows that the AIC is minimized at the discrete number $L = 5$.

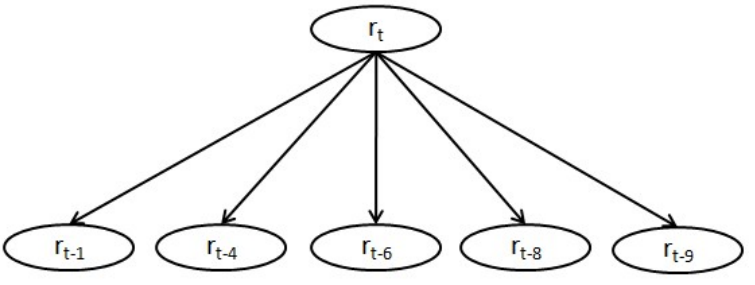


Fig. 4 Bayesian network

5.2 Network Determination

The discrete value sets at $L = 5$ are indicated in Table 2. The parameter $c_l(r^l)$ is the discrete value and the parameters $(C_l)_{\min}$ and $(C_l)_{\max}$ denote the maximum and minimum values among the samples in the cluster C_l. The Bayesian network, which is determined from the data of Table 2, is shown in Fig. 4. The number of samples means the number of stock price returns included in the cluster. The notation $c_l(r^l)$ denotes the cluster center; the discrete value. We notice that the return r_t has the dependency on the one-day prior return r_{t-1}, the 4-days prior return r_{t-4}, the 6-days prior return r_{t-6}, the 8-days prior return r_{t-8}, and the 9-days prior return r_{t-9}.

5.3 Comparison of Prediction Accuracy

The prediction accuracy is compared in Table 3. The row label AR(9), MA(6), ARMA(9,6) and ARCH(9,9) denote the results by AR model with $p = 9$, MA model with $q = 6$, ARMA model with $p = 9$ and $q = 6$ and ARCH model with $p = 9$ and

Table 2 Discrete value set of TOYOTA Motor Corp. stock return

Cluster	$(C_l)_{\min}$, $(C_l)_{\max}$	$c_l(r^l)$
C_1	[-21.146%, -3.900%)	-5.55%
C_2	[-3.900%, -1.285%)	-2.10%
C_3	[-1.285%, 0.000%]	-0.47%
C_4	(0.000%, 2.530%]	1.08%
C_5	(2.530%, 16.264%]	4.12%

Table 3 Comparison of prediction accuracy

	Max. error	Min. error	Ave. error	Correlation coefficient
AR(9)	7.5091	0.0615	0.7448	2.6657
MA(6)	7.6259	0.0319	0.7417	2.6859
ARMA(9,6)	7.1204	0.0401	0.7427	2.6739
ARCH(9,9)	8.0839	0.0597	0.7527	2.6992
Bayesian Network	7.8579	0.0415	0.7122	3.1494

$q = 9$, respectively. Besides, the row label "Bayesian network" means the result by the present method.

Table 3 shows that, in the present algorithm, the average error is about 5%-smaller than the others and the correlation coefficient is improved by 10 to 20% against the others.

6 Conclusions

The stock price return prediction algorithm using Bayesian network was presented in this study. The stock price return distribution is discretized by the Ward method. Bayesian network models the stochastic dependency of the stock price returns. The present algorithm was compared with the traditional time-series prediction algorithms such as AR, MA, ARMA and ARCH models in the TOYOTA motor corporation stock price prediction. The present algorithm show, in the average error, about 5% better than the time-series prediction algorithms.

References

1. Box, G., Jenkins, G.M., Reinsel, G.C.: Time Series Analysis: Forecasting and Control. Prentice-Hall (1994)
2. Brockwell, P.J., Davis, R.A.: Time Series: Theory and Methods. Springer (2009)
3. Mills, T.C.: Time Series Techniques for Economists. Cambridge University Press (1990)
4. Box, G., Jenkins, G.M.: Time Series Analysis: Forecasting and Control, 2nd edn., Oakland (1976)
5. Engle, R.F., Ng, V.K.: Measuring and testing the impact of news on volatility. Journal of Finance 48(5), 1749–1778 (1993)
6. Bollerslevb, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
7. Nelson, D.B.: Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347–370 (1991)
8. Takayasu, H.: Practical Fruits of Econophysics. Springer (2006)
9. Pearl, J., Russell, S.: Bayesian networks. In: Arbib, M.A. (ed.) Handbook of Brain Theory and Neural Networks, pp. 157–160. MIT Press (2002)
10. Ben-Gal, I.: Bayesian networks. In: Ruggeri, F., Kennett, R.S., Faltin, F.W. (eds.) Encyclopedia of Statistics in Quality and Reliability, vol. 1, John Wiley & Sons (2007)
11. Akaike, H.: Information theory and an extention of the maximum likelihood principle. In: Proc. 2nd International Symposium on Information Theory, pp. 267–281 (1973)

12. Heckerman, D., Geiger, D., Chickering, D.: Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning 20(3), 197–243 (1995)
13. Cooper, G.F., Herskovits, E.: A Bayesian method for the induction of probabilistic networks from data. Machine Learning 9(4), 309–347 (1992)
14. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Representation and Reasoning Series. Morgan Kaufmann (1988)
15. Ward, J.H.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58, 236–244 (1963)

Building a Decision Support Tool for Taiwan's National Health Insurance Data – An Application to Fractures Study

Chien-Lung Chan, Phan Dinh Van, and Nan-PingYang

Abstract. This study aims to construct a decision support tool for handling the Taiwan's National Health Insurance data from the researchers' perspectives. The National Health Research Institute (NHRI) provides data for researchers to conduct health related research by applications. Since the provided insurance data is in unstructured text, it is necessary to transfer those data into a data warehouse or data mart such as SQL Server. However, the data retrieval and synthesis are still very complicated through a number of complex SQL queries to export the final figures and tables. This causes a major problem for doctors, medical and insurance researchers who are not familiar with SQL Server. Therefore, this study proposes the tools to extract and save the data in SQL Server 2008, which facilitates the data synthesis. The tools not only support doctors, medical and insurance researchers to obtain the final analysis for nationwide epidemiology but also help them to make good decisions in building health care strategies.

Keywords: Decision support system, Data warehouse, Taiwan's National Health Insurance Data, Data transfer, SQL Server.

Chien-Lung Chan
Department of Information Management, College of Informatics, Yuan Ze University, Taoyuan, Taiwan

Phan Dinh Van
Community Health Research Center & Institute of Public Health, National Yang-Ming University, Taipei, Taiwan

Nan-PingYang
Department of Medical Research & Department of Orthopedic Surgery, Taoyuan General Hospital, Department of Health, Taoyuan, Taiwan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 407–417.
springerlink.com

1 Introduction

The accumulated large amount of data in many fields has created the urgent needs of data storage, data retrieval and data analysis. With the development of information technology, the Internet, especially, the theories about database, scientists have managed to solve the problem above by proposing the data warehouse to save data and to facilitate data retrieval (Rainardi 2008). Analytical processing that involves very complex queries and few or no updates usually termed decision support is one of the primary uses of data warehouses, hence the terms data warehousing and decision support often are found together (Widom 1995).

In Taiwan, there is a huge data warehouse to record all patients' healthcare insurance information called National Health Insurance Research Database (NHIRD), which is also provided to the researchers to conduct health-related studies (Malinowski et al. 2008). The data extracted for distribution is in unstructured text format. Nevertheless, almost the researches need to synthesize the data in certain criteria according to their needs. To do that, the data should be re-organized in structured data so that it would be convenient to connect with each other and easily to be extracted and analyzed.

This study addressed the national health insurance data and its disadvantages that we have experienced as we are the project team members on the analysis of Taiwan's National Health Insurance database (Yang et al., 2010, 2011, and Lee et al., 2011). We realized the urgent need of extracting and re-organizing the original data into a data warehouse to facilitate the data retrieval for future research by using effective and suitable tools. Therefore, this study attempts to investigate the following questions and develop such tools at the same time:

How to build the data warehouse for the data extracted from NHIRD?
How can data be retrieved and synthesized for future research?
What kind of decision support tools can be constructed?
What kind of decisions can be supported by these tools?

2 Literature Review

A data warehouse is a collection of subject-oriented, integrated, nonvolatile, and time-varying data to support management decisions (Malinowski et al. 2008). Apart from the definition of data warehouse, other definitions related to our study such as, cube, dimension, OLAP (Online Analytical Processing) are provided.

OLAP is the activity of interactively analyzing business transaction data stored in the dimensional data warehouse to make tactical and strategic business decisions. Typical people who do OLAP work are business analysts, business managers, and executives (Rainardi 2008). However, it also can be used in the healthcare fields to view the patients' data. OLAP support models for forecasting, trends analyzing and statistical analysis. It is operated based on two constructs: cube and dimensions.

Data cube is a popular model used to conceptualize the data in a data warehouse. The data cube contains points or "cells" that are measures or values based on a set of dimensions. Dimensions often form a hierarchy and allow different levels of granularity in the warehouse (Datta and Thomas 1999).

The National Health Insurance data is divided into different categories (Table 1). Hence, it is necessary to distinguish the type and year of the data when transferring it to SQL Server.

Each table has many attributes based on the different structure of data for different category, such as CD data, DD data, OO data, and so on. The structure of each category data was available in the National Health Insurance Research Database Coding Book. After transferring and changing all the non-structured data into the structured data in SQL Server, the data can be retrieved by using the support tools provided by SQL Server.

Table 1 Catergories of Taiwan's National Health Insurance Database.

Category	Description
HOSB	Registry for contracted medical facilities
HOSX	Supplementary registry for contracted medical facilities
HOX	Registry for medical services
BED	Registry for contracted beds
DOC	Registry for board-certified specialists
HV	Registry for catastrophic illness patients
CT	Monthly claim summary for ambulatory care claims
CD	Ambulatory care expenditures by visits
OO	Details of ambulatory care orders
DD	Inpatient expenditures by admissions
DO	Details of inpatient orders
GD	Expenditures for prescriptions dispensed at contracted pharmacies
DRUG	Registry for drug prescriptions
ID	Registry for beneficiaries
PER	Registry for medical personnel
DETA	Registry for contracted specialty services
GO	Details of prescriptions dispensed at contracted pharmacies
DT	Monthly claim summary for inpatient claims

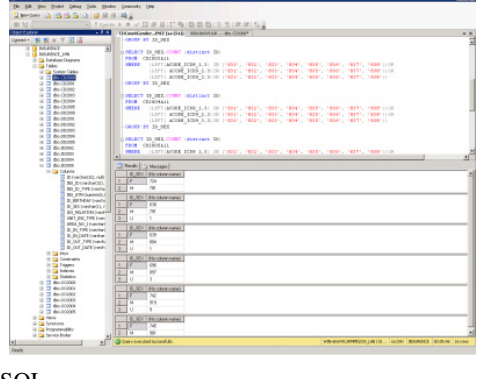


Fig. 1 Retrieve data by SQL

3 Deployment of Decision Support Tools

3.1 Overview of Decision Support Tools

The target users for decision support tool are those novices in database. It does not only perform the simple retrievals but also perform the complicated retrieval and data synthesis, draw charts and export into other data, such as: SQL Server, MS-Excel data and structural text data. They are used to retrieve many requirements automatically.

The NHI data is transferred to the data warehouse via data transfer tool, and then the requirements of data retrieval will be implemented through the decision support tool as shown in Fig. 2. Data warehouse is organized on database management system SQL Server 2008. Since the NHI data are very large, retrieval time and calculation of results can be very slow.

3.2 Building Data Transfer Tool

Then NHI data is stored as non-structured attributes and each paragraph contains all the information of the attributes in each record. Therefore, when the data is transferred into the warehouse, it is necessary to separate the attributes based on the structure of each data category as shown in Fig. 3.

The structured file is a MS-Excel file used to determine the attribute name, data type, number of characters and the order of initial character. It must be loaded before the data is transferred, because the program base on mentioned structure to cut and transfer data.

3.3 Building Decision Support Tool

Decision support tool will support users to retrieve the data from the data warehouse easily. Results are illustrated in tables, charts, or in other different data. This tool provides a form for users to specify requests, then, the program will automatically operate each request.

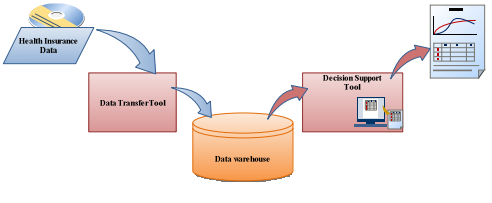


Fig. 2 Model of data warehouse and tools

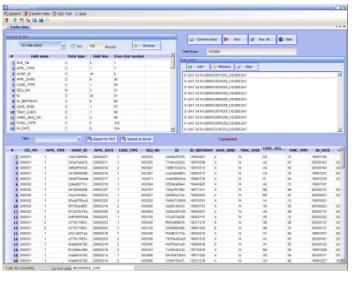


Fig. 3 Main form of data transfer tool

3.4 A Retrieval form of Decision Support Tool

We designed this tool in some various kinds of interfaces in order to separate requirements of the users. It is used to perform general queries on each table. Decision support tool is used to retrieval data depending on different dimensions such as the time, value of attributes, geographic distribution, population, and multi-dimensional, etc. Moreover, these tools also have other functions, such as automatic retrieval data, export a schema into a new one, union some schema to one, creating a new database, backing up data, restoring data and data cleansing as shown in Fig. 4.

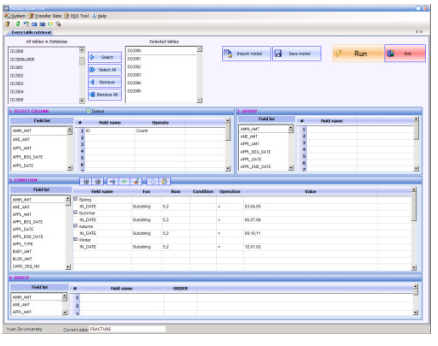


Fig. 4 A form of decision support tool

4 Application of NHIRD in Fractures Study

After finishing development of the tools, we applied them in analyzing cases of fracture from 2000 to 2005 based on the data extracted from the NHIRD. To support the data interpretation, we use two concepts: (1) subjects and (2) number of doctor visits. Subject is patient defined by attribute "ID" in database; each value of ID represents one subject. Each subject may see the doctor many times. The number of doctor visits is calculated by the frequency of the doctor visiting.

4.1 Data Protection and Permission

Data in the NHIRD that could be used to identify patients or care providers, including medical institutions and physicians, is scrambled before being sent to the NHRI for database inclusion, and is further scrambled before being released to each researcher. Theoretically, it is impossible to query the data alone to identify individuals at any level using this database. All researchers who wish to use the NHIRD and its data subsets are required to sign a written agreement declaring that they have no intention of attempting to obtain information that could potentially violate the privacy of patients or care providers. This study protocol was evaluated by the NHRI, who gave their agreement to the planned analysis of the NHIRD (Application and Agreement Number: 98018).

4.2 Analyzing Different Kinds of Fracture Subjects

Fig. 5 below shows numbers of subjects in 17 different fractures from 2000 to 2005 given one million sampling population in Taiwan. In general, almost all kinds of subject numbers of fractures increase by years. Among them, Spine fracture and fracture of Ulnar & Radius are higher than 15 others. Spine fracture's tendency grows faster than fracture of Ulnar & Radius from 2002 to 2005, with the number of Spine fracture the highest. Next, we can see that the fracture of Tibia & Fibula's trend and fracture of Phalanges bone's trend also increase by year, but fracture of Tibia & Fibula decrease from 2004 to 2005. So, we can propose some warning for the patient groups who can possibly face with the risk of those four kinds of fractures.

4.3 Number of Doctor Visits for Fracture

To understand the healthcare resources utilization, we can look at Fig. 6, which shows the number of doctor visits for different kinds of fracture. By observing the fractures of Ulnar & Radius, Spine and Tibia & Fibula carefully, we found these three kinds of fracture all tend to increase from 2000 to 2004. Ulnar & Radius and Spine fracture increase much more significantly than Tibia & Fibula fracture. The number of doctor visits is highest for Ulnar & Radius fracture. Nonetheless, since 2004, fracture of the Ulnar & Radius and Tibia & Fibula decrease slightly.

4.4 Effects of Months on Fracture Subjects

The number of fracture subjects in December is the highest from 2000 to 2005 and it has a clear distance with other months. The remaining eleven months were quite the same but grew up in 6 years. As shown in Fig. 7, all these results suggest that we need to warn people to be especially careful in December. Meanwhile, people should be informed that December is the time with the highest risk in fracture. In addition, we may examine events, festivals or weather in December to see their correlation with fracture.

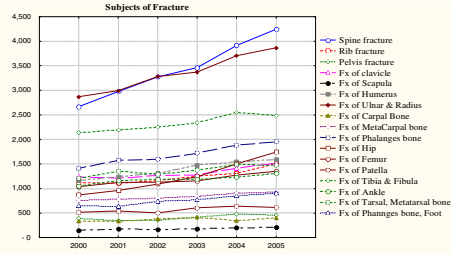


Fig. 5 Trend of fracture subjects 2000 - 2005

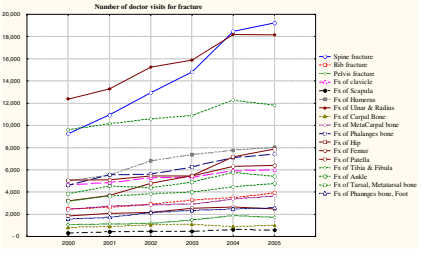


Fig. 6 Trends of doctor visiting times for fractures 2000-2005

4.5 Effects of Seasons on Fracture Subjects

Number of fracture subjects in winter can be easily distinguished with other seasons. The number of fractures in all seasons has regularly risen from 2000 to 2005 as shown in Fig.8. Among them, winter is the highest one with 6,615 subjects in 2000 to 10,145 in 2005, an average annual increase rate of 8.95%. The other three seasons has also grown slighter, with the average of 8.46% in spring, 8.34% in summer and 7.63% in autumn.

From this result, we can realize that winter has affected fractures more than the other seasons. It seems probable that raining, slippery walking and other reasons are causes of fracture. Therefore, we should create awareness and educate people to be extra careful during winter. For example, through TV programs, warning notices, and so on.

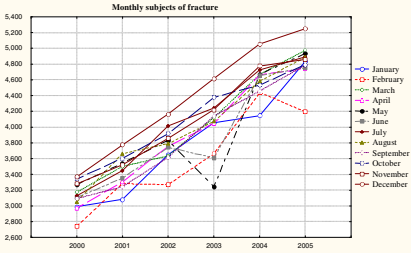


Fig. 7 Trends of monthly fracture subjects.

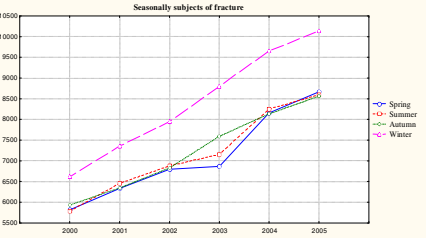


Fig. 8 Trends of seasonally fracture subjects.

4.6 Effects of Age on Fractures

We divided patients into 4 groups: (1) below or equal to 20 years old, (2) 20 to 39 years old, (3) 40 to 59 years old and (4) 60 years old and above. Fractures in all age groups increased from 2000 to 2005. As shown in Fig. 9, among them, the 60 years old and above group is the highest with the average increasing rate of 7.22% per year. Subsequently, 40-50 years old group has the average of 2.85% a year, 20-39 years old group is 1.75% and the last group, below 20 years old is 0.35%. The elderly face with the highest risk in fracture. Thus, we should implement policies to look after them and prevent them from fracture.

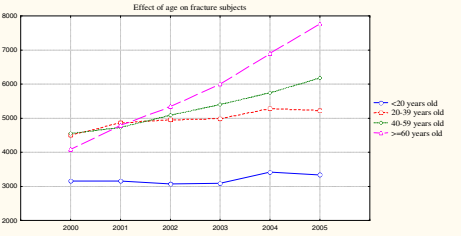


Fig. 9 Effects of age on fracture subjects.

4.7 Effects of Age and Gender on Fractures

By cross table analysis of gender and age group, we found the male / female ratio of fracture patients who are less than or equal to 59 years old are always higher than 1 with the average of 1.97 in less than 20 years old group, average of 1.91 in 20-39 years old group and average of 1.06 in 40-59 years old group; whereas the 60 years old and above group is less than 1 with the average of 0.53. It means that the fracture subjects of female are almost two times of male for the group equal to or more than 60 years old. Male fracture subjects are almost two times of female before 40 years old.

5 Discussion and Conclusion

Our study has successfully addressed all the research questions. First, we built a data warehouse to process the Taiwan's National Health Insurance data and it can help in many ways for future research work. In the beginning, the data were in unstructured text and had to be processed before it can be used. Next, we developed a Data transfer tool to convert and transfer data into the data warehouse. This tool can help us transfer data from unstructured text to the data warehouse easily. Moreover, the Decision support tool was built to retrieve data for various research works. This tool can execute many complex queries to obtain results for research related to national healthcare insurance. Finally, we applied those tools for the study of fractures based on the one million sampling data from 2000 to 2005 extracted from the NHIRD. In the application of fracture study, we analyzed the trends of various kinds of fractures, healthcare resources utilization of fracture, and influencing factors such as month, season, age, gender, temperature and precipitation. Based on the above results, health promotion decision makers can make sound decisions to prevent fracture events. Among all kinds of fractures, it is essential to pay special attention on fracture of spine, Ulnar & Radius, Tibia & Fibula, and Phalanges bone. Decision makers also need to consider the seasonal effect especially winter particularly in December. The occurrence of fracture events more frequent during winter compared with other seasons or months. Strategic decisions are needed to prevent fracture events in the society. Additionally, training programs can be developed and warning notices can remind people to practice caution in this time.

Acknowledgement. The research was supported by a grant from National Science Council, Taiwan (NSC 99-2221-E-155-045-MY3 & NSC 100-2632-S-155-001).

References

1. Datta, A., Thomas, H.: The cube data model: a conceptual model and algebra for on-line analytical processing in data warehouses. Decision Support Systems 27(3), 289–301 (1999)

2. Lee, Y.H., Hsu, Y.N., Yu, I.L., Phan, D.V., Chou, P., Chan, C.L., Yang, N.P.: Treatment incidence of and medical utilization for hospitalized subjects with pathologic fractures in Taiwan-Survey of the 2008 National Health Insurance Data. BMC Health Services Research 11(1), 230–236 (2011)
3. Lin, Y.J., Tian, W.H.: Urbanization and the utilization of outpatient services under National. Health Policy 103, 236–243 (2011)
4. Malinowski, E., Zimányi, E., et al.: Advanced Data Warehouse Design: From conventional to spatial and temporal applications. Springer, Heidelberg (2008)
5. Rainardi, V.: Building a Data Warehouse: With Examples in SQL Server. Springer, New York (2008)
6. Widom, J.: Research Problems in Data Warehousing. In: CIKM Conference on Information and Knowledge Management, pp. 25–30. ACM, Baltimore (1995)
7. Yang, N.P., Chan, C.L., Yu, I.L., Lee, C.Y., Chou, P.: Estimated prevalence of orthopaedic fractures in Taiwan-Across-sectional study based on nationwide insurance data. Injury, Int. J. Care Injured 41(12), 1266–1272 (2010)
8. Yang, N.P., Chen, H.C., Phan, D.V., Yu, I.L., Chan, C.L., Chou Renn, J.H.: Epidemiological survey of orthopedic joint dislocations based on nationwide insurance data in Taiwan, 2000-2005. BMC Musculoskeletal Disorders 12(1), 253–260 (2011)

Appendices

Table A1 Number of fracture subjects 2000-2005.

ICD9 Code	FRACTURE	2000	2001	2002	2003	2004	2005	Mean	S.D.
805.X-806.X	Spine fracture	2,669	2,975	3,276	3,461	3,914	4,243	3,423.00	584.08
807.X	Rib fracture	1,122	1,107	1,166	1,250	1,307	1,515	1,244.50	153.09
808.X	Pelvis fracture	388	345	362	417	479	460	408.50	53.50
810.X	Fracture of clavicle	1,243	1,207	1,268	1,271	1,410	1,519	1,319.67	119.59
811.X	Fracture of Scapula	145	175	172	182	206	210	181.67	23.99
812.X	Fracture of Humerus	1,167	1,248	1,310	1,479	1,548	1,586	1,389.67	171.81
813.X	Fracture of Ulnar & Radius	2,870	2,993	3,279	3,371	3,704	3,865	3,347.00	388.28
814.X	Fracture of Carpal Bone	341	338	378	410	347	400	369.00	31.47
815.X	Fracture of Metacarpal bone	757	783	810	841	908	931	838.33	69.18
816.X	Fracture of Phalanges bone	1,410	1,575	1,599	1,720	1,884	1,954	1,690.33	204.03

Table A1 (*continued*)

820.X	Fracture of Hip	869	961	1,092	1,239	1,502	1,745	1,234.67	334.93
821.X	Fracture of Femur	1,040	1,122	1,128	1,156	1,274	1,353	1,178.83	114.02
822.X	Fracture of Patella	516	542	508	609	640	614	571.50	56.37
823.X	Fracture of Tibia & Fibula	2,143	2,194	2,248	2,344	2,550	2,481	2,326.67	162.18
824.X	Fracture of Ankle	1,049	1,168	1,188	1,185	1,231	1,306	1,187.83	84.26
825.X	Fracture of Tarsal, Metatarsal bone	1,212	1,356	1,293	1,372	1,487	1,477	1,366.17	105.96
826.X	Fracture of Phalanges bone, Foot	648	641	736	770	855	910	760.00	108.54

Table A2 Numbers of fracture subjects for each month 2000-2005.

Month	2000	2001	2002	2003	2004	2005	Mean	S.D.
January	2,989	3,083	3,647	4,055	4,146	4,835	3,792.50	700.12
February	2,740	3,285	3,273	3,660	4,440	4,196	3,599.00	634.06
March	3,172	3,503	3,635	4,125	4,667	4,975	4,012.83	703.70
April	2,972	3,303	3,760	4,050	4,665	4,745	3,915.83	715.18
May	3,269	3,541	3,828	3,241	4,667	4,936	3,913.67	724.84
June	3,109	3,348	3,743	3,608	4,659	4,754	3,870.17	684.16
July	3,130	3,449	4,014	4,243	4,726	4,893	4,075.83	694.23
August	3,051	3,664	3,793	4,075	4,582	4,883	4,008.00	660.30
September	3,096	3,240	3,618	4,143	4,462	4,783	3,890.33	680.52
October	3,342	3,604	3,920	4,377	4,532	4,794	4,094.83	565.30
November	3,281	3,521	3,855	4,220	4,775	4,860	4,085.33	649.98
December	3,369	3,775	4,165	4,617	5,056	5,250	4,372.00	735.40

Table A3 Seasonally subjects of fracture.

Season	2000	2001	2002	2003	2004	2005	Mean	S.D.
Spring	5,818	6,333	6,800	6,870	8,155	8,671	7,107.83	1,091.37
Summer	5,782	6,457	6,886	7,157	8,246	8,593	7,186.83	1,067.40
Autumn	5,933	6,342	6,829	7,595	8,135	8,563	7,232.83	1,035.44
Winter	6,615	7,358	7,948	8,800	9,654	10,145	8,420.00	1,360.25

Table A4 Effect of age on fracture subjects.

Age Stratum	2000	2001	2002	2003	2004	2005	Mean	S.D.
<20 years old	3,163	3,159	3,076	3,088	3,424	3,341	3,208.50	141.84
20-39 years old	4,503	4,869	4,955	4,989	5,288	5,217	4,970.17	279.65
40-59 years old	4,549	4,725	5,090	5,407	5,741	6,174	5,281.00	617.47
>=60 years old	4,092	4,799	5,343	6,007	6,890	7,762	5,815.50	1,356.99

Table A5 Gender rate (male / female) in different age groups.

Age Stratum	2000	2001	2002	2003	2004	2005	Mean	S.D.
<20 years old	1.70	1.93	1.93	1.99	2.12	2.15	1.97	0.16
20-39 years old	1.91	2.04	1.86	1.93	1.86	1.86	1.91	0.07
40-59 years old	1.04	1.01	1.03	1.05	1.07	1.12	1.06	0.04
>=60 years old	0.56	0.55	0.53	0.52	0.52	0.51	0.53	0.02

Case of Process Mining from Business Execution Log Data

Joonsoo Bae and Young Ki Kang

Abstract. Process Mining is one research area to find useful information from various processes in business execution log data. As BPMS, ERP, SCM, etc so called process recognizing information system are widely spread, process mining research are getting emphasis recently. Also execution log data of one enterprise are frequently utilized for checking current status of management or analysis of resource efficiency. But since analysis result depends on measurement criteria and method, it is very important to select systematic process mining algorithm based on business model or enterprise strategy. In this paper many process mining techniques are introduced and compared using process mining tool, ProM.

Keywords: Process mining, Log data.

1 Introduction

As information technology develops and enterprise management environment changes, the business processes of enterprise are changing and complex continuously. Therefore, each enterprise makes effort to improve business processes efficiently and renovate them in order to increase competiveness. Business Process Management(BPM) is a holistic management approach focused on aligning all aspects of an organization with the wants and needs of clients. It promotes business effectiveness and efficiency while striving for innovation, flexibility, and integration with technology. BPM attempts to improve processes continuously. (OMG, 2008)

Especially Business Process Reengineering (BPR) is still regarded as critical concept. In order to improve productivity by BPR, the matter to be settled first is defining current business process correctly. But when workflow management system is introduced to design and support business process, it is very difficult to

Joonsoo Bae, · Young Ki Kang
BPM Lab, Chonbuk National University, South Korea
e-mail: {jsbae,kykyou}@jbnu.ac.kr

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 419–425.
springerlink.com

design complete business process from the scratch. The reason is that since workflow system developer cannot understand business process completely, they can not consider all of the business properties. The execution log data was provided for BPR from the system level in the past, but the monitoring tool was basic level that provides only fragment data of ongoing process using simple database field.

For this business process designing, process mining attracts attention nowadays. Process Mining is for extracting information from event logs. For example, the audit trails of a workflow management system or the transaction logs of an enterprise resource planning system can be used to discover models describing processes, organizations, and products.(Aalst et al. 2004)

2 Application Case Researches

This section shows how process mining techniques are applied to the real enterprise from the execution log data. After the give log data are filtered by some criteria, necessary process mining techniques are chosen. The chosen process mining technique can be used to check current process, analyze, and select future process. The improvement can be made by using process modification, such as node modification, arc modification and clustering. The modified process is also analyzed.

2.1 Log Data Filtering

The real execution log data of telecommunication service can be summarized as in Table 1. The period of the data is from March 20 to March 22, 2009.(Jung et al. 2010)

Table 1 Summary of execution log data

Event	49,998
Cases	10,933
Activities(Event class)	39
Resources(Originators)	12
Attributes	2
Timestamp	Type:yyyy-mm-dd-yyhh:mm:ss
Time Duration	Start: 2009-03-20 23:26:14 End : 2009-03-22 20:14:19

The log data can be extracted in the form of CSV or Excel file. The process mining tool ProM do not allow CSV format, it should be converted into XML file.(http://www.processminig.org)

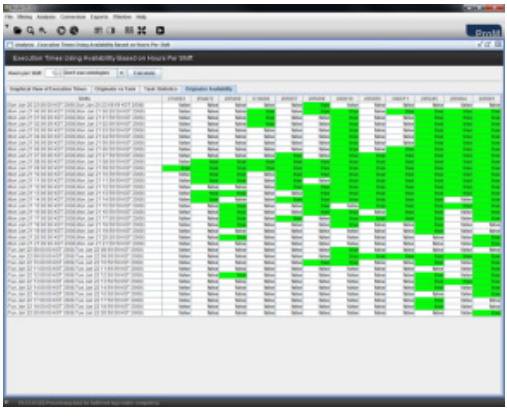


Fig. 1 Sample data of execution log at one branch

2.2 Log Data Analysis

The converted XML file should be analyzed for correctness before process mining. In Figure 1, theaverage execution time per hour can be viewed by Execution times Using Availability Based on Hours Per Shift. Currently Task 205061(OO branch) is occupied all day.

The average execution time is in Figure 2, KR05(Reception of Fault Report during Breakdown) has the longest average execution time, 362 minutes.

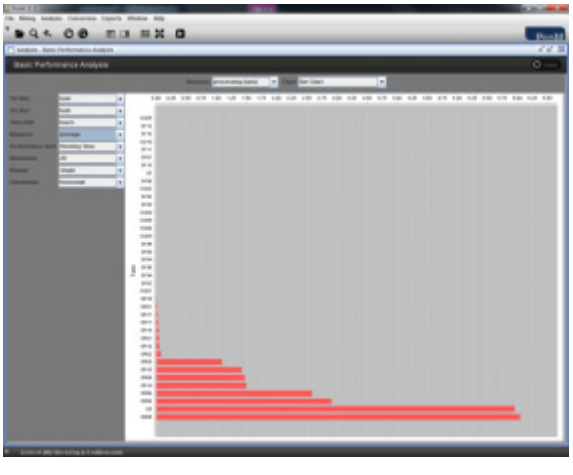


Fig. 2 Average Execution Time

2.3 Process Mining

The visible process model is needed in order to check how task is performed, this technique is Fuzzy mining. The result of Fuzzy mining is in Figure 3. The green graph in the bottom shows conformance level, 78.5%, which is relatively low because this is initial mining using average value between node and arc.(Jung et al. 2009)

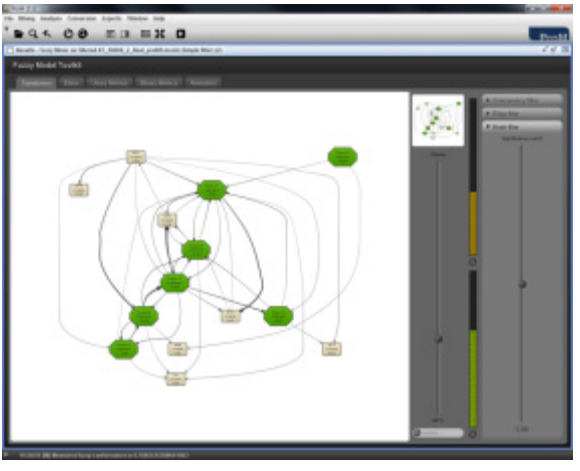


Fig. 3 Fuzzy Mining

Figure 4 shows the result of low edge deletion. By increasing edge value as 0.252, it means removing all edges below 0.252 signal value. This is reasonable because if the two nodes (tasks) have low signal value, then they are not so related. So the edge between two nodes can be removed.

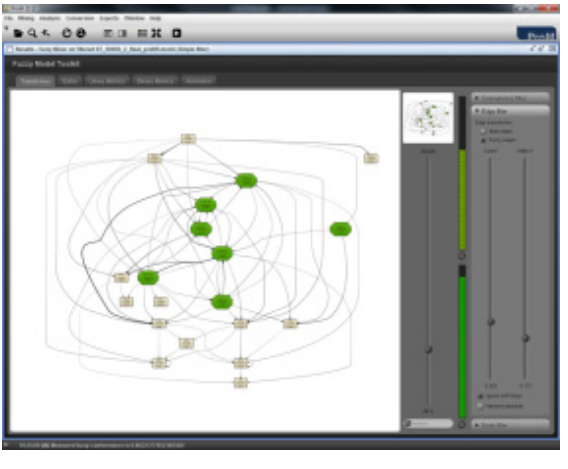


Fig. 4 Deletion of low edge

In the figure, there are 7 clustered nodes in green hexagon. The detail internal of the clustered node can be seen and checked how the internal nodes are connected to the external nodes. The clustered nodes are log signal values, which can be used to organizational integration or movement.

Table 2 Cluster Information and Connection

Cluster Number	Connected Nodes
48	KR02-KF02-KQ02
43	KR17-KQ05
45	KR12-KQ03-KR16-KQ04
42	KR09-KR51-KR10
44	KQ08-KQ10-KF10
47	KR71-KF01

Next step is to increase conformance level between nodes. First, Trace conformance inspector checks disconnected edges. We can see that KQ01 and KR14 in Case 2008012100024501 are disconnected. KQ01-KR14 is that after ClientInformationQuery(KQ01), the problem was resolved (KR14) during consulting. This case can happen, so the edge should be connected. Using edit toolkit in the Fuzzy mining, KQ01 and KR14 are connected. This will make conformance level to 97.78%. The advantage of this method shows we can get more benefit (higher conformance level) by using partial clustering and filtering rather than disconnecting arcs and nodes.

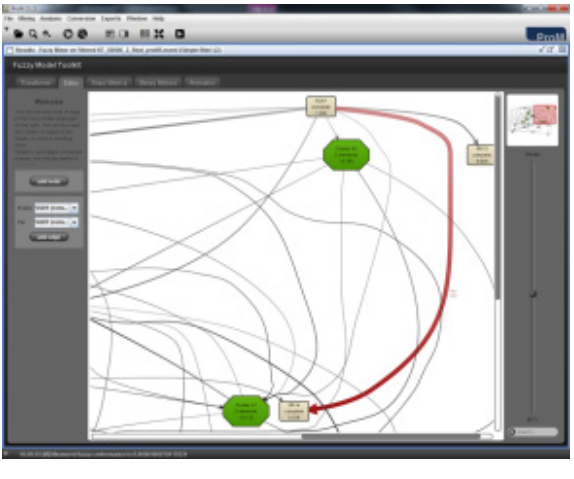


Fig. 5 Adding connection between two nodes

By using Fuzzy mining simulation, task flow of KQ01(Client Information Query) and finish time of individual tasks can be identified as in Figure 6.

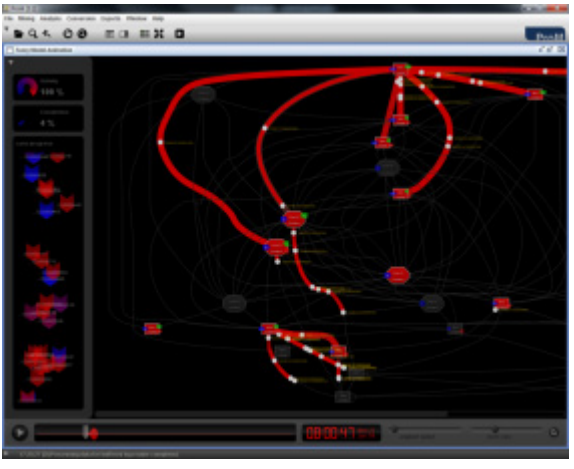


Fig. 6 Fuzzy mining simulation

2.4 Process Analysis

The sequence of Fuzzy mining is summarized in Table 3.There are 5 steps, from the initial analysis to last step (Node and Arc Editing) the performance criteria, log conformance increase consistently.

Table 3 Fuzzy Mining Sequence and Result

Analysis Sequence	Filter Node	Filter Edge	Log conformance	Model Detail
Initial Analysis	0.500	0.200	**78.85%**	40.50 %
Expansion to the maximum node arc	0.06	1.0	**92.68%**	100%
Node Clustering	0.356	1.0	**94.44%**	72%
Best Arc Selection	0.356	0.252	**96.22%**	72%
Node and Arc Editing	0.356	0.252	**97.78%**	72%

3 Conclusion and Future Research

The main content of this paper is process mining case of real service execution log data. The total sequence of process mining is: Log data filtering, Log data analysis, and process analysis. Especially Fuzzy mining technique was applied in this case. Through this process mining, one process design was extracted from execution log data, and the procedure how the conformance level can be improved is provided. In the future research, we can develop many kinds of process mining algorithms and show how we can select best mining technique from analysis results.

Acknowledgement. This research was supported by Basic Science Research Program through theNational Research Foundation of Korea(NRF) funded by the Ministry ofEducation, Science and Technology(2010-0025650).

References

Object Management Group/Business Process Management Initiative, BPMN 1.1 : OMG Specification (February 2008)

van der Aalst, Weijters, A.J.M.M.: Process Mining: research agenda. Comput. Ind. 53(3) (2004)

http://www.processmining.org/online/

Jung, J.-Y., Bae, J., Liu, L.: Hierarchical Clustering of Business Process Models. International Journal of Innovative Computing, Information and Control 5(12(A)), 4501–4511 (2009)

Jung, J.-Y., Bae, J.: Rule-Based Ad-Hoc Workflow Modeling for Service Coordination: A Case Study of a Telecom Operational Support System. IEICE Trans.

Consistent Conjectures in Mixed Oligopoly with Discontinuous Demand Function

Nataliya I. Kalashnykova, Vyacheslav V. Kalashnikov,
and Mario A. Ovando Montantes

Abstract. In this paper, a conjectured variations equilibrium state (CVE) in a mixed oligopoly model is studied. The agents make conjectures concerning the variations of the clearing price in a dependence upon variations in their production volumes. Existence and uniqueness theorems are established for the conjectured variations equilibrium (called an exterior equilibrium) for any set of feasible conjectures. To introduce the notion of an interior equilibrium, the authors develop a consistency criterion for the conjectures (referred to as influence coefficients). Next, an existence theorem for the interior equilibrium (understood as a CVE with consistent conjectures) is proved.

1 Introduction

Since recently, papers and monographs dealing with behavioral patterns of agents of mixed markets have become very common (*see*, e.g., [12], [14] and references

Nataliya I. Kalashnykova
Graduate School of Mathematical Sciencs, FCFM, UANL, Av. Universidad S/N,
San Nicolás de los Garza, N.L., Mexico 66450; and Sumy State University,
Rimsky-Korsakov st. 2, Sumy, Ukraine 44007
e-mail: nkalash2009@gmail.com

Vyacheslav V. Kalashnikov
ITESM, Campus Monterrey, Av. Eugenio Garza Sada 2501 Sur, Monterrey, N.L.,
Mexico 64849; and Central Economics and Mathematics Institute (CEMI),
Russian Academy of Sciences (RAS), Nakhimovsky pr. 47, Moscow, Russia 117418;
and Sumy State University, Rimsky-Korsakov st. 2, Sumy, Ukraine 44007
e-mail: kalash@itesm.mx

Mario A. Ovando Montantes
ITESM, Campus Monterrey, Av. Eugenio Garza Sada 2501 Sur, Monterrey, N.L.,
Mexico 64849
e-mail: naven_15@hotmail.com

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 427–436.
springerlink.com

therein). A homogeneous commodity (oligopoly) market is called *mixed* whenever a state-owned (public, domestic, etc.) welfare-maximizing agent (company) competes against profit-maximizing private (foreign) firms. The models studied mainly differ in the definition of the objective (utility) function maximized by the public firm. Previous papers by the authors [12], [14] examined the mixed oligopoly where the public agent was interested in increasing domestic social surplus, which is almost uniformly accepted in the literature in the following form:

$$S(p;q_0,q_1,\ldots,q_n) = \int_0^{\sum_{i=0}^{n} q_i} p(x)\,dx - p\cdot\left(\sum_{i=1}^{n} q_i\right) - b_0 - \frac{1}{2}a_0 q_0^2. \qquad (1)$$

(the notation will be introduced and explained in the next section).

However, there also exists another concept of the special agent called a *labor-managed* firm, whose utility function is represented with the company's income per worker. The seminal paper by Ward [25] is considered as a "first swallow". Since then, many researchers have investigated the equilibrium in such a mixed oligopoly, *cf.* [19], [8], [23], [24], [4], [6], [3], [7], [16], [20], [17], [9], [21], [5], to mention only few. Excellent surveys of other works dealing with labor-managed companies can be also found in [1] and [22].

The *conjectural variations equilibrium* (CVE) as another possible solution concept in static games. According to this concept, players behave in the following manner: each player chooses her/his most favorable action taking into account that every rival's strategy is a *conjectured function* of her/his own strategy.

In their previous works [10], [13], [15], the authors examined a conjectural variations oligopoly model, in which the degree of influence on the whole situation by each agent is modeled by special parameters (influence coefficients). There, a new gamut of conjectural variations equilibria (CVE) was introduced and investigated, in which the conjectural variations (represented through the influence coefficients of each agent) affected the structure of the Nash equilibrium. In other words, not only a classical Cournot competition (where the influence coefficients are all 1) but also a Cournot-type model with influence coefficient values different from 1 were considered. Various equilibrium existence and uniqueness results were obtained in the above-cited works.

The consistency (or, sometimes, "rationality") of the equilibrium is defined as the coincidence between the conjectural best response of each agent and the conjectured reaction function of the same. A conceptual difficulty arises when one considers consistency in the case of many agents (*see* [12]). To cope with this conceptual difficulty arising in many players models, Bulavsky [2] proposed a completely new approach. It is supposed that each player makes conjectures *not* about the (optimal) response functions of the other players but only about the *variations of the market price* depending upon his infinitesimal output variations. Knowing the rivals' conjectures (called *influence coefficients*), each agent can realize certain verification procedure and check out if his/her influence coefficient is *consistent* with the

others'. Exactly the same verification formulas were obtained independently by Liu et al. in [18] who established the existence and uniqueness of a consistent conjectural variation equilibrium in an electricity market.

In their previous papers [11]–[12], the authors extended the results obtained in [2] to a *mixed* oligopoly model with a continuously differentiable objective function. In [14], the restriction is relaxed and the demand function is allowed, similar to [2], to have a finite number of points, at which not only its derivative but the function itself may suffer a break. Nevertheless, the same results previously obtained in [11]–[12] were proved to hold under these, more general conditions.

In this work, we extend results obtained in [14] to the case of oligopoly with a labor-managed company competing with profit-maximizing agents. The paper is organized as follows. In Section 2, we describe the mixed oligopoly mathematical model from with a labor-managed agent in case of not necessarily smooth demand function. Then, in Section 3, we define the concept of an *exterior equilibrium*, i.e., a conjectural variations equilibrium (CVE) with the influence coefficients fixed in an exogenous manner. The existence and uniqueness theorem for this kind of CVE ends the section. Section 4 deals with a further advanced concept of an *interior equilibrium*, which is defined as an exterior equilibrium with *consistent* conjectures (influence coefficients). The consistency criterion, verification procedure, and the existence theorem for the interior equilibrium are formulated in the same Section 4. Acknowledgements and the reference list complete the paper.

2 Model Specification

Consider a market of a homogeneous good with al least 3 suppliers/producers having the cost functions $f_i = f_i(q_i)$, $i = 0, 1, \ldots, n$, $n \geq 1$, where $q_i \geq 0$ is the output volume by supplier i. Consumers' demand is described by a demand function $G = G(p)$, whose argument p is the market price proposed by the producers. An active demand value D is non-negative and independent of price. Since the demand function is allowed to have points of discontinuity (breaks), the equilibrium between the demand and supply for a given price p is described by the following (balance) inequalities

$$g(p) + D \leq \sum_{i=0}^{n} q_i \leq G(p) + D. \tag{2}$$

Here, and from now on, we denote the right limit of the function G at (any) point p by $g(p) = G(p+0) = \lim\limits_{\zeta \to p+0} G(\zeta)$ while assuming that the left limit of this function at (any) point p coincides with its proper value: $G(p) = G(p-0) = \lim\limits_{\zeta \to p-0} G(\zeta)$.

We accept the following assumptions about the model's data.

A1. The demand function $G = G(p) \geq 0$ is defined on the domain $p \in (0, +\infty)$ being non-increasing and continuously differentiable everywhere except for a finite number of points $0 < p_1 < p_2 < \ldots < p_K$, at which both the function and its

derivative suffer breaks. Therefore, one has $g(p) = G(p)$ at the continuity points, whereas $g(p) < G(p)$ at the break-points.

A2. For each $i = 1,\dots,n$, the cost function $f_i = f_i(q_i)$ is twice continuously differentiable with $f_i'(0) > 0$ and $f_i''(q_i) > 0$ for all $q_i \geq 0$ (that is, f_i is strictly convex over R_+, for any $i = 1,2,\dots,n$.) As for $i = 0$, we assume that cost function has the form

$$f_0(q_0) = (\bar{a}q_0 + \bar{b})\bar{f}_0(q_0), \tag{3}$$

where $\bar{a} > 0, \bar{b} > 0$, and the function $\bar{f}_0$ is twice continuously differentiable with $\bar{f}_0'(0) > 0$ and $\bar{f}_0''(q_0) > 0$ for all $q_0 \geq 0$. (The latter of course simply means that the function $\bar{f}_0$ is stricty convex over R_+).

Private producer/supplier $i = 1,\dots,n$, chooses her/his output volume $q_i \geq 0$ so as to maximize her/his profit function $\pi_i(p,q_i) = p \cdot q_i - f_i(q_i)$. On the other hand, the labor-managed company $i = 0$ selects its supply/production value $q_0 \geq 0$ with an aim to maximize the company's income per worker

$$W(p;q_0,q_1,\dots,q_n) = \frac{p \cdot q_0 - f_0(q_0)}{l(q_0)} = p \cdot \frac{q_0}{\bar{a}q_0 + \bar{b}} - \bar{f}_0(q_0), \tag{4}$$

where $l(q_0) = \bar{a}q_0 + \bar{b}$ is the labor input function. This assumption means that the marginal quantity of labor used is constant.

Now we postulate that the agents (both public and private) assume that their choice of production volumes may affect the price value p. The latter assumption could be defined by a *conjectured* dependence of the price p upon the output volumes q_i. If so, the first order maximum condition to describe the equilibrium would have the form: for the public company (with $i = 0$)

$$\frac{\partial W}{\partial q_0} = p\frac{\bar{b}}{\left(\bar{a}q_0 + \bar{b}\right)^2} + \frac{q_0}{\bar{a}q_0 + \bar{b}}\frac{\partial p}{\partial q_0} - \bar{f}_0'(q_0)\begin{cases} = 0, & \text{if } q_0 > 0; \\ \leq 0, & \text{if } q_0 = 0; \end{cases} \tag{5}$$

and

$$\frac{\partial \pi_i}{\partial q_i} = p + q_i\frac{\partial p}{\partial q_i} - f_i'(q_i)\begin{cases} = 0, & \text{if } q_i > 0; \\ \leq 0, & \text{if } q_i = 0, \end{cases} \quad \text{for} \quad i = 1,2,\dots,n. \tag{6}$$

Thus, one can see that to describe the agent's behavior one need evaluate the behavior of the derivative $\dfrac{\partial p}{\partial q_i} = -v_i$ rather than the complete dependence of p upon q_i. We introduce the minus here in order to deal with non-negative values of v_i. Of course, the conjectured dependence of p on q_i must provide (at least local) concavity of the i-th agent's conjectured profit as a function of his/her output. Otherwise, one cannot guarantee the profit to be *maximized* (but not *minimized*). As we suppose that the cost functions f_i are quadratic and hence convex, then, for $i = 1,\dots,n$, the concavity of the product $p \cdot q_i$ with respect to the variation η_i of the current production volume will be enough. For instance, it is sufficient to assume the coefficient v_i (from now on, referred to as the i-th agent's *influence coefficient*) to be

non-negative and constant. Then, for the private companies $i = 1,\dots,n$, the conjectured local dependence of the profit upon the production output η_i has the form $[p - v_i(\eta_i - q_i)]\,\eta_i - f_i(\eta_i)$, while the maximum condition at $\eta_i = q_i$ is provided by the relationships

$$\begin{cases} p = v_i q_i + f_i'(q_i), & \text{if } q_i > 0; \\ p \le f_i'(0), & \text{if } q_i = 0. \end{cases} \tag{7}$$

Similarly, the labor-managed company conjectures the local dependence of domestic social surplus on its production output q_0 in the form

$$\frac{[p - v_0(\eta_0 - q_0)]\,\eta_0 - f_0(\eta_0)}{l(\eta_0)}, \tag{8}$$

which allows one to write down the maximum condition at $\eta_0 = q_0$ as follows:

$$\begin{cases} p = v_0 \dfrac{(\bar{a}q_0 + \bar{b})\,q_0}{\bar{b}} + \dfrac{(\bar{a}q_0 + \bar{b})^2}{\bar{b}} \bar{f}_0'(q_0), & \text{if } q_0 > 0; \\ p \le \bar{b}\bar{f}_0'(0), & \text{if } q_0 = 0. \end{cases} \tag{9}$$

Were the agents' conjectures about the model given exogenously like it was assumed in [11]–[12], we would allow the values v_i to be functions of q_i and p. However, here we use the approach from [2], where the conjecture parameters v_i for an equilibrium state are determined simultaneously with the price p and the output values q_i by a special verification procedure. In the latter case, the influence coefficients are the scalar parameters determined only for the equilibrium. In Section 4, such an equilibrium state is referred to as *interior* one and is described by the set of variables and parameters $(p, q_0, q_1, \dots, q_n, v_0, v_1, \dots, v_n)$.

3 Exterior Equilibrium

In order to introduce a verification procedure we need a notion of equilibrium called *exterior* (*cf.*, [2], [11]–[12]) with parameters v_i fixed (given exogenously).

Definition 1. A vector $(p, q_0, q_1, \dots, q_n)$ is called an *exterior equilibrium* for given influence coefficients $(v_0, v_1, \dots, v_n)$, if the market is balanced, i.e., condition (2) is satisfied, and for each i the (first order) maximum conditions (7) and (9) are valid.

In what follows, we are going to consider only the case when the set of really producing participants is fixed (i.e., it does not depend upon the values v_i of the influence coefficients). To provide for this, we make the following assumption.

A3. For the price value $p_0 = \max\left\{\max\limits_{1\le j\le n} b_j, \bar{f}_0'(0)\bar{b}\right\}$, for any $i = 1, 2, \dots, n$, there exists (uniquely, due to **A2**) a (positive) output volume q_i^0 such that $p_0 = f_i'(q_i^0)$. Moreover, for $i = 0$, there exists (uniquely, due to **A2**) a (positive) output volume q_0^0 such that:

$$p_0 = \frac{(\bar{a}q_0^0 + \bar{b})^2}{\bar{b}} \bar{f}_0'(q_0^0). \tag{10}$$

Finally, it is assumed that

$$\sum_{i=0}^{n} q_i^0 < G(p_0). \tag{11}$$

It is straightforward that assumptions **A1**–**A3** guarantee that for all non-negative values of $v_i, i = 0,1,\ldots,n$, and there always exists a unique solution of the optimality conditions (7) and (9) satisfying the balance equality (2), i.e., an exterior equilibrium. Moreover, conditions (2), (7) and (9) can hold simultaneously if, and only if $p > p_0 \geq \max\limits_{0\leq i\leq n} f_i'(0)$, that is, if and only if all outputs q_i are strictly positive, $i = 0,1,\ldots,n$.

Theorem 1. *Under assumptions A1, A2 and A3, for any $D \geq 0$, $v_i \geq 0$, $i = 1,\ldots,n$, and $v_0 \in [0,\bar{v}_0)$, there exists uniquely an exterior equilibrium $(p,q_0,q_1,\ldots,q_n)$, which depends continuously upon the parameters $(D,v_0,v_1,\ldots,v_n)$. The equilibrium price $p = p(D,v_0,v_1,\ldots,v_n)$ is always greater than p_0 (i.e., $p(D,v_0,v_1,\ldots,v_n) > p_0$), and, as a function of these parameters, is continuous with respect to both D and v_i, $i = 0,1,\ldots,n$. Moreover,*

$$\frac{\partial p}{\partial D}(D+0) = \begin{cases} \dfrac{1}{\sum\limits_{i=1}^{n} \dfrac{1}{v_i + f_i''(q_i)} + [\bar{q}_0(p,v_0)]' - G'(p+0)}, & \text{if} \quad \sum\limits_{i=0}^{n} q_i = g(p) + D; \\ 0, & \text{if} \quad \sum\limits_{i=0}^{n} q_i > g(p) + D; \end{cases} \tag{12}$$

$$\frac{\partial p}{\partial D}(D-0) = \begin{cases} \dfrac{1}{\sum\limits_{i=1}^{n} \dfrac{1}{v_i + f_i''(q_i)} + [\bar{q}_0(p,v_0)]' - G'(p-0)}, & \text{if} \quad \sum\limits_{i=0}^{n} q_i = G(p) + D; \\ 0, & \text{if} \quad \sum\limits_{i=0}^{n} q_i < G(p) + D. \end{cases} \tag{13}$$

Here, $[\bar{q}_0(p,v_0)]'$ is the partial derivative (with respect to p) of the function of solution of optimality conditions (5), which is well-defined on the non-negative orthant $(p,v_0) \in R_+^2$.

Remark 1. Having connected with a vertical segment the extreme points $(p_k, G(p_k))$ and $(p_k, g(p_k))$ of the gap corresponding to a discontinuity point $p = p_k$, $k = 1,2,\ldots,K$, and thus having closed the (plane) graph $\Gamma = \{(p,G(p))\}_{p\in(0,+\infty)}$ of the function G, one gets a curve L called the *demand curve*. All points of this curve satisfy the inequalities $g(p) \leq G \leq G(p)$. Now let us (formally) define the one-side derivatives of the function G at the "kink" points of the curve L as follows: $G'(p_k \pm 0) = -\infty$, $k = 1,2,\ldots,K$. An exterior equilibrium $(p,q_0,q_1,\ldots,q_n)$ defines a point $(p,G) \in L$ such that $G = \sum\limits_{i=0}^{n} + D$. Based on the latter notation and definitions, formulas (12)–(13) can be rewritten uniformly as the following one:

$$p'(D\pm 0)\equiv\frac{\partial p}{\partial D}(D\pm 0)=\frac{1}{\sum\limits_{i=1}^{n}\dfrac{1}{v_i+f_i''(q_i)}+[\bar{q}_0(p,v_0)]'-G'(p\pm 0)}. \tag{14}$$

Therefore, the inequality $p'(D+0)\neq p'(D-0)$ can happen *only* at an equilibrium state associated with one of the curve L's "kink" point. Moreover, if $p'(D+0)<p'(D-0)$ then it is a *concavity* point, whereas the inequality $p'(D+0)>p'(D-0)$ implies it to be a *convexity* point.

4 Interior Equilibrium

Now we are ready to define an interior equilibrium. To do that, we first describe the procedure of verification of the influence coefficients v_i as it was given in [2], [11]–[12]. Consider an exterior equilibrium $(p,q_0,q_1,\ldots,q_n)$ existing for some $(v_0,v_1,\ldots,v_n)$ and D. Assume that one of suppliers, say k, temporarily changes his/her behavior by abstaining from maximization of the conjectured profit (or domestic social surplus, as it is in case $k=0$) and producing infinitesimal variations around his/her equilibrium output volume q_k. In mathematical terms, it is tantamount to restricting the model agents to the subset $\{i:\ i\neq k\}$ with the output q_k subtracted from the active demand.

Variation of the production output by agent k is then equivalent to the (corresponding) active demand variation in form $D_k=D-q_k$. We can assume that by observing the corresponding (infinitesimal) variations of the equilibrium price, agent k gets the derivative of the equilibrium price with respect to the active demand, i.e., his/her influence coefficient. Indeed, applying formulas (12) or(13) from Theorem 1 (or, which is the same, universal relationships (14) from Remark 1) to calculate the derivatives, one has to remember that agent k has been temporarily excluded from the equilibrium model, hence the term with number $i=k$ has to be deleted from all the sums. Keeping that in mind, we come to the following criterion.

4.1 Consistency Criterion

At an exterior equilibrium $(p,q_0,q_1,\ldots,q_n)$, the influence coefficients v_k, $k=0,1,\ldots,n$, are referred to as *consistent* if there exist values r_k, $k=0,1,\ldots,n$, such that

$$r_k\in\begin{cases}[G'(p+0),G'(p-0)], & \text{if}\quad G'(p+0)\le G'(p-0);\\ [G'(p-0),G'(p+0)], & \text{otherwise;}\end{cases} \tag{15}$$

and the following equalities hold:

$$v_0=\frac{1}{\sum\limits_{j=1}^{n}\dfrac{1}{v_j+f_j''(q_i)}-r_0}, \tag{16}$$

and

$$v_k = \frac{1}{\sum\limits_{j=1,j\neq k}^{n} \frac{1}{v_j + f_j''(q_j)} + [\bar{q}_0(p, v_0)]' - r_k}, \; k = 1, \ldots, n. \tag{17}$$

Remark 2. It is easy to check that in the particular case of $n = 2$ (a mixed *duopoly* model), formulas (16) and (17) become considerably simpler:

$$v_0 = \frac{1}{\frac{1}{v_1 + f_1''(q_1)} - r_0}, \tag{18}$$

and

$$v_1 = \frac{1}{[\bar{q}_0(p, v_0)]' - r_1}, \tag{19}$$

respectively.

Now we are in a position to define the concept of an interior equilibrium.

Definition 2. A collection $(p; q_0, q_1, \ldots, q_n; v_0, v_1, \ldots, v_n)$, where $p \geq 0, q_k \geq 0$, $v_k \geq 0$, $k = 0, 1, \ldots, n$, is referred to as an *interior equilibrium* if, for the given influence coefficients $v_0, v_1, \ldots, v_n$, the collection $(p; q_0, q_1, \ldots, q_n)$ is an exterior equilibrium, and the consistency criterion (16)–(17) is satisfied for all $k = 0, 1, \ldots, n$. If, in addition, all the values of parameters r_k, $k = 0, 1, \ldots, n$, are the same, i.e., $r_0 = r_1 = \cdots = r_n = r(\leq 0)$, then the interior equilibrium is called *strong*. Here, the one-side derivatives $G'(p-0)$ and $G'(p+0)$ are calculated at the corresponding points of the demand curve L with the components $\left(p, G = \sum\limits_{i=0}^{n} q_i + D\right)$. If a coefficient $r_k = -\infty$, then the corresponding $v_k = 0$.

Remark 3. If all agents are profit-maximizing private companies, then formulas (16)–(17) reduce to the uniform ones published independently by Bulavsky [2] in 1997 and by Liu et al. [18] in 2007 (however, the latter authors considered only smooth demand functions, so they actually defined strong interior equilibrium with all $r_k = r$, $k = 1, \ldots, n$):

$$v_k = \frac{1}{\sum\limits_{j=1,j\neq k}^{n} \frac{1}{v_j + f_j''(q_j)} - r_k}, \; k = 1, \ldots, n. \tag{20}$$

The following theorem is an extension of Theorem 2 from [2] to the case of a mixed oligopoly.

Theorem 2. *Under assumptions A1, A2, and A3, there exists an interior equilibrium.*

In order to apply the concept of consistent conjectured variations equilibrium (CCVE) to investigation of the dynamics of demand at a market of a homogeneous

commodity, we need to study the existence and behavior of the (strongy) consistent influence coefficients $v_k = v_k(r)$ as functions of the parameter r. Recall that this parameter equals $G'(p)$ at a continuity point of the demand function G and belongs to the interval with the limits $G'(p-0)$ and $G'(p+0)$ when p is a discontinuity point. These results will be obtained in our future papers.

5 Conclusions

In this paper, we consider a model of mixed (with at least one labor-managed agent being present) oligopoly with conjectural variations equilibrium (CVE) and not necessarily continuous demand function. The agents' conjectures concern the price variations depending upon their production output's increase or decrease. We establish existence and uniqueness results for the conjectured variations equilibrium (called an exterior equilibrium) for any set of feasible conjectures. To introduce the notion of an interior equilibrium, we develop a consistency criterion for the conjectures (referred to as influence coefficients) and prove the existence theorem for the interior equilibrium (understood as a CVE with consistent conjectures).

In order to prepare the base for the study of dynamics in demand and supply, in our future research work, we will also investigate the behavior of the consistent conjectures in dependence upon a parameter representing the demand function's (sub-)derivative with respect to the market price.

Appendix

The first author was supported by the PAICYT project No. CE250-09 and by the SEP-CONACYT project CB-2009-01-127691. The research activity of the second and third authors was financially supported by the R&D Department (Cátedra de Investigación) CAT-174 of the Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), Campus Monterrey, and by the SEP-CONACYT projects CB-2008-01-106664, Mexico.

References

1. Bonin, J.P., Putterman, L.: Economics of Cooperation and the Labor-managed Economy. Harwood Academic Publisher, Chur (1987)
2. Bulavsky, V.A.: Structure of demand and equilibrium in a model of oligopoly. Economics and Mathematical Methods (Ekonomika i Matematicheskie Metody) 33, 112–134 (1997)
3. Chiarella, C.: The dynamic operation of the market for partnership deeds in a duopolistic industry of a profit-maximizing firm and workers' enterprise. Managerial and Decision Econ. 14, 451–458 (1993)
4. Cremer, H., Crémer, J.: Duopoly with employee-controlled and profit-maximizing firms: Bertrand vs Cournot competition. J. Comparative Econ. 16, 241–258 (1992)
5. Cuccia, T., Cellini, R.: Workers' enterprises and the taste for production: the arts, sport and other cases. Scottish J. Political Econ. 56, 123–137 (2009)

6. Delbono, F., Rossini, G.: Competition policy vs horizontal merger with public, entrepreneurial, and labor-managed firms. J. Comparative Econ. 16, 226–240 (1992)
7. Futagami, K., Okamura, M.: Strategic investment: the labor-managed firm and the profit-maximizing firm. J. Comparative Econ. 23, 73–91 (1996)
8. Horowitz, I.: On the effects of Cournot rivalry between entrepreneurial and cooperative firms. J. Comparative Econ. 15, 115–121 (1991)
9. Ireland, N.J.: Random pricing by labor-managed firms in markets with imperfect consumer information. J. Comparative Econ. 31, 573–583 (2003)
10. Isac, G., Bulavsky, V.A., Kalashnikov, V.V.: Complementarity, Equilibrium, Efficiency and Economics. Kluwer Academic Publishers, Dordrecht (2002)
11. Kalashnikov, V.V., Bulavsky, V.A., Kalashnykova, N.I.: How to compute consistent conjectures in a mixed oligopoly. In: Clute, R.C. (ed.) Proceedings of the 10th Annual International Economics & Business Research Conference (IBER & TLC 2010), pp. 1–11 (2010)
12. Kalashnikov, V.V., Bulavsky, V.A., Kalashnykova, N.I., Castillo Pérez, F.J.: Mixed oligopoly with consistent conjectures. European J. Oper. Res. 210, 729–735 (2011)
13. Kalashnikov, V.V., Cordero, E., Kalashnikov Jr., V.V.: Cournot and Stackelberg equilibrium in mixed duopoly models. Optimization 59, 689–706 (2010)
14. Kalashnikov, V.B., Kalashnykova, N.I., Ovando Montantes, M.A., Kalashnikov Jr., V.V.: A mixed oligopoly model with a discontinuous demand function. In: Clute, R.C. (ed.) Proceedings of the 2011 Las Vegas International Academic Conference, pp. 867–876 (2011)
15. Kalashnikov, V.V., Kemfert, C., Kalashnikov Jr., V.V.: Conjectural variations equilibrium in a mixed duopoly. European J. Oper. Res. 192, 717–729 (2009)
16. Lambertini, L.: Cournot vs Stackelberg equilibria with entrepreneurial and labour managed firms. J. Internat. Comparative Econ. 5, 31–43 (1997)
17. Lambertini, L., Rossini, G.: Capital commitment and Cournot competition with labour-managed and profit-maximising firms. Australian Econ. Papers 37, 14–21 (1998)
18. Liu, Y.F., Ni, Y.X., Wu, F.F., Cai, B.: Existence and uniqueness of consistent conjectural variation equilibrium in electricity markets. Int. J. Electrical Power Energy Sys. 29, 455–461 (2007)
19. Mai, C., Hwang, H.: Export subsidies and oligopolistic rivalry between labor-managed and capitalist economies. J. Comparative Econ. 13, 473–480 (1989)
20. Neary, H.M., Ulph, D.: Strategic investment and the co-existence of labour-managed and profit-maximising firms. Canadian J. Econ. 30, 308–328 (1997)
21. Ohnishi, K.: Strategic investment in a new mixed market with labor-managed and profit-maximizing firms. Metroeconomica 59, 594–607 (2008)
22. Putterman, L.: Labour-managed firms. In: Durlauf, S.N., Blume, L.E. (eds.) The New Palgrave Dictionary of Economics, vol. 4, pp. 791–795. Palgrave Macmillan, Basingstoke (2008)
23. Stewart, G.: Strategic entry interactions involving profit-maximising and labour-managed firms. Oxford Economic Papers 43, 570–583 (1991)
24. Stewart, G.: Management objectives and strategic interactions among capitalist and labour-managed firms. J. Econ. Behavior and Organization 17, 423–431 (1992)
25. Ward, B.: The firm in Illyria: Market syndicalism. Amer. Econ. Rev. 48, 566–589 (1958)

Convergence Properties of Criteria Weights of Dominant AHP with Solving Equations of Interchangeability

Takafumi Mizuno, Shin Sugiura, and Eizo Kinoshita

Abstract. This paper focus on convergence properties of interchangeable criteria weight vectors of Dominant AHP(Analytic Hierarchy Process). We find a degree of freedom of the vectors by solving equations which express interchangeability among them. We also reconsider a relation between interchangeability and the law of CRU (Constant of Ratio in Unit Evaluation), an optimality of solutions of CCM (Concurrent Convergence Method), and a relation between Dominant AHP and ANP(Analytic Network Process).

1 Introduction, Interchangeability of Dominant AHP

Dominant AHP(Analytic Hierarchy Process) [1] [2] is a decision making process which selects one from plural alternatives by scoring them for an ultimate purpose.

Let us consider there are m alternatives and n criteria, and let $m \times n$ matrix A is an *evaluation matrix*, d-th alternative is *dominant*, i-th alternative is *controlling*, and a *criteria weight vector* $\mathbf{b}^d$ which corresponds to the dominant alternative. The Dominant AHP's scoring procedure calculates *aggregate scores* $\mathbf{e}^{i:d}$ as:

$$\mathbf{e}^{i:d} = \frac{(AA_i^{-1})(A_iA_d^{-1})\mathbf{b}^d}{\mathbf{1}^\top(AA_i^{-1})(A_iA_d^{-1})\mathbf{b}^d} = \frac{AA_d^{-1}\mathbf{b}^d}{\mathbf{1}^\top AA_d^{-1}\mathbf{b}^d}. \tag{1}$$

Takafumi Mizuno
Meijo University, Kani, Japan
e-mail: tmizuno@urban.meijo-u.ac.jp

Shin Sugiura
Meijo University, Kani, Japan
e-mail: shinsu@urban.meijo-u.ac.jp

Eizo Kinoshita
Meijo University, Kani, Japan
e-mail: kinoshit@urban.meijo-u.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 437–445.
springerlink.com

Under a fixed dominant alternative, if the evaluation matrix A is rigid then aggregate scores depend only on the dominant alternative whichever alternative we choose as controlling; we simply write the scores as $\mathbf{e}^d$ when d-th alternative is dominant. In this paper, we focus on finding criteria weight vectors which have next property:

Definition 1. Interchangeability We say that criteria weight vectors $\mathbf{b}^1, \mathbf{b}^2, \cdots, \mathbf{b}^m$ are *interchangeable* when $\mathbf{e}^1 = \mathbf{e}^2 = \cdots = \mathbf{e}^m$.

We introduce notations for algebraic analysis used throughout this paper. Symbol $^\top$ is a transpose operator. $\mathbb{R}$ is the set of real numbers. $\mathbb{R}_+ = \{r | r > 0, r \in \mathbb{R}\}$. $\mathbb{Q}_n = \{\mathbf{x} | \mathbf{1}_n^\top \mathbf{x} = 1, \mathbf{x} \in \mathbb{R}_+^n\}$. $\mathbf{1}_n^\top = [1, .., 1]$ is a n-dimensional vector which consists of only 1, and is used for normalization; $\mathbf{x} \in \mathbb{R}_+^n$ is normalized as $\frac{\mathbf{x}}{\mathbf{1}_n^\top \mathbf{x}}$. We omit n and write simply $\mathbf{1}^\top$ without confusion. Notice that $\mathbf{1}^\top \mathbf{x}$ is scalar. m is number of alternatives, or number of rows of A, and n is number of criteria, or number of columns of A. Each element a_{ij} of A is in $\mathbb{R}_+$, and each vector $\mathbf{a}_j$ consists of elements of j-th columns of A. A_k is $n \times n$ diagonal square matrix, its main diagonal elements are elements of k-th row of A, and outsides the diagonal are zero.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}, \quad \mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad A_k = \begin{bmatrix} a_{k1} & & & \\ & \ddots & & \\ 0 & & 0 & \\ & & & a_{kn} \end{bmatrix} \tag{2}$$

2 Solve Equations of Interchangeable Criteria Weight Vectors

In this section, we derive a main theorem of this paper. The theorem is:

Theorem 1. *Interchangeable criteria weight vectors be expressed with* $(m-1)(n - \mathrm{rank}(A)) + n$ *parameters. We obtain the vectors by normalizing linear combinations of the parameters.*

In other words, interchangeable criteria weight vectors exist infinitely and its degree of freedom is $(m-1)(n-\mathrm{rank}(A))+n$. If $\mathrm{rank}(A) = n$ then the degree equals n.

In processes of Dominant AHP, nonlinearity of normalizations complicates finding interchangeable criteria weight vectors. A next lemma, however, says that we have only to consider without normalizations: a set of unnormalized vectors which are interchangeable without normalizations, if we normalize each vector of the set, equals a set of normalized vectors which are interchangeable with normalizations.

Lemma 1. *Let* $V = \{(\mathbf{B}^1, \mathbf{B}^2, \cdots, \mathbf{B}^m) | AA_1^{-1}\mathbf{B}^1 = AA_2^{-1}\mathbf{B}^2 = \cdots = AA_m^{-1}\mathbf{B}^m, \quad \mathbf{B}^1, \cdots, \mathbf{B}^m \in \mathbb{R}_+^n\}$, $\dot{V} = \{(\frac{\mathbf{B}^1}{\mathbf{1}^\top \mathbf{B}^1}, \cdots, \frac{\mathbf{B}^m}{\mathbf{1}^\top \mathbf{B}^m}) | (\mathbf{B}^1, \cdots, \mathbf{B}^m) \in V\}$, $v = \{(\mathbf{b}^1, \mathbf{b}^2, \cdots, \mathbf{b}^m) | \frac{AA_1^{-1}\mathbf{b}^1}{\mathbf{1}^\top AA_1^{-1}\mathbf{b}^1} = \frac{AA_2^{-1}\mathbf{b}^2}{\mathbf{1}^\top AA_2^{-1}\mathbf{b}^2} = \cdots = \frac{AA_m^{-1}\mathbf{b}^m}{\mathbf{1}^\top AA_m^{-1}\mathbf{b}^m}, \quad \mathbf{b}^1, \mathbf{b}^2, \cdots, \mathbf{b}^m \in \mathbb{Q}_n\}$. *Then* $\dot{V} = v$.

Proof. Let $(\frac{\mathbf{B}^1}{\mathbf{1}^\top\mathbf{B}^1},\cdots\frac{\mathbf{B}^m}{\mathbf{1}^\top\mathbf{B}^m})\in\dot{V}$, each $\frac{\mathbf{B}^i}{\mathbf{1}^\top\mathbf{B}^i}$ is in $\mathbb{Q}_n$ and $\frac{AA_i^{-1}\frac{\mathbf{B}^i}{\mathbf{1}^\top\mathbf{B}^i}}{\mathbf{1}^\top AA_i^{-1}\frac{\mathbf{B}^i}{\mathbf{1}^\top\mathbf{B}^i}}=\frac{AA_i^{-1}\mathbf{B}^i}{\mathbf{1}^\top AA_i^{-1}\mathbf{B}^i}$.
While $AA_1^{-1}\mathbf{B}^1=\cdots=AA_m^{-1}\mathbf{B}^m$, then $(\frac{\mathbf{B}^1}{\mathbf{1}^\top\mathbf{B}^1},\cdots,\frac{\mathbf{B}^m}{\mathbf{1}^\top\mathbf{B}^m})$ is in v ; $\dot{V}\subseteq v$.
Let $v'=\{(\frac{\mathbf{b}^1}{\mathbf{1}^\top AA_1^{-1}\mathbf{b}^1},\cdots,\frac{\mathbf{b}^m}{\mathbf{1}^\top AA_m^{-1}\mathbf{b}^m})|(\mathbf{b}^1,\cdots,\mathbf{b}^m)\in v\}$. Then $v'\subset V$. ($\because$ substitute $\frac{\mathbf{b}^i}{\mathbf{1}^\top AA_i^{-1}\mathbf{b}^i}$ for $\mathbf{B}^i$ in the definition of V.)
We introduce two functions T and S.
$T:v\ni(\mathbf{x}^1,\cdots,\mathbf{x}^m)\mapsto(\tau_1(\mathbf{x}^1),\cdots,\tau_m(\mathbf{x}^m))$, each $\tau_i(\mathbf{x})=\frac{\mathbf{x}}{\mathbf{1}^\top AA_i^{-1}\mathbf{x}}$.
And $S:v'\ni(\mathbf{y}^1,\cdots,\mathbf{y}^m)\mapsto(\sigma(\mathbf{y}^1),\cdots,\sigma(\mathbf{y}^m))$, $\sigma(\mathbf{y})=\frac{\mathbf{y}}{\mathbf{1}^\top\mathbf{y}}$.
$T:v\to v'$, and $S:v'\to\dot{V}$ ($\because v'\subset V$) ; $S\cdot T:v\to\dot{V}$.
If $\sigma(\tau_i(\mathbf{x}))=\sigma(\tau_i(\mathbf{s}))$,$\mathbf{x},\mathbf{s}\in\mathbb{Q}_n$ then $\mathbf{x}=\mathbf{s}$. ($\because\sigma(\frac{\mathbf{x}}{\mathbf{1}^\top AA_i^{-1}\mathbf{x}})=\sigma(\frac{\mathbf{s}}{\mathbf{1}^\top AA_i^{-1}\mathbf{s}})$, $\frac{\frac{\mathbf{x}}{\mathbf{1}^\top AA_i^{-1}\mathbf{x}}}{\mathbf{1}^\top\frac{\mathbf{x}}{\mathbf{1}^\top AA_i^{-1}\mathbf{x}}}=\frac{\frac{\mathbf{s}}{\mathbf{1}^\top AA_i^{-1}\mathbf{s}}}{\mathbf{1}^\top\frac{\mathbf{s}}{\mathbf{1}^\top AA_i^{-1}\mathbf{s}}}$, $\frac{\mathbf{x}}{\mathbf{1}^\top\mathbf{x}}=\frac{\mathbf{s}}{\mathbf{1}^\top\mathbf{s}}$, and $\mathbf{1}^\top\mathbf{x}=\mathbf{1}^\top\mathbf{s}=1$.)
This means that if $\mathbf{x}\in\mathbb{Q}_n$ then $\sigma(\tau_i(\mathbf{x}))$ is injective; $S\cdot T$ is injective.
$v\supseteq\dot{V}$, and v has an injective function into $\dot{V}$, the function is $S\cdot T$, thus $\dot{V}=v$. □

Let $\mathbf{B}^d$ is a criteria weight vector without normalization when d-th alternative is dominant; $\mathbf{B}^d\in\mathbb{R}_+^n$. And B_j^d is j-th element of $\mathbf{B}^d$. Let $\mathbf{E}^d$ is an aggregate score without normalization, and E_i^d is i-th element of $\mathbf{E}^d$. $\mathbf{E}^d$ is calculated as:

$$\mathbf{E}^d=AA_d^{-1}\mathbf{B}^d,\quad\text{each element is } E_i^d=\sum_{j=1}^{n}\frac{a_{ij}}{a_{dj}}B_j^d. \tag{3}$$

Interchangeability is expressed as:

$$\mathbf{E}^d=\mathbf{E}^{d+1},\quad d=1,\cdots,m-1. \tag{4}$$

$$AA_d^{-1}\mathbf{B}^d-AA_{d+1}^{-1}\mathbf{B}^{d+1}=0. \tag{5}$$

This is expressed in an equation system which consists of $m(m-1)$ equations:

$$\mathscr{A}\mathbf{w}=\begin{bmatrix}\mathscr{A}_1 & -\mathscr{A}_2 & & & \\ & \mathscr{A}_2 & -\mathscr{A}_3 & & 0 \\ & & \ddots & & \\ & 0 & & \mathscr{A}_{m-2} & -\mathscr{A}_{m-1} \\ & & & & \mathscr{A}_{m-1} & -\mathscr{A}_m\end{bmatrix}\mathbf{w}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}, \tag{6}$$

where $\mathbf{w}=\left[B_1^1\cdots B_n^1B_1^2\cdots B_n^2\cdots\cdots B_1^m\cdots B_n^m\right]^\top$, and

$$\mathscr{A}_d=AA_d^{-1}=\begin{bmatrix}\frac{a_{11}}{a_{d1}} & \frac{a_{12}}{a_{d2}} & \cdots & \frac{a_{1n}}{a_{dn}}\\ \frac{a_{21}}{a_{d1}} & \frac{a_{22}}{a_{d2}} & \cdots & \frac{a_{2n}}{a_{dn}}\\ \vdots & & \ddots & \vdots\\ \frac{a_{m1}}{a_{d1}} & \frac{a_{m2}}{a_{d2}} & \cdots & \frac{a_{mn}}{a_{dn}}\end{bmatrix}=\left[\frac{1}{a_{d1}}\mathbf{a}_1\ \cdots\ \frac{1}{a_{dn}}\mathbf{a}_n\right]. \tag{7}$$

The column size of the matrix $\mathscr{A}$ is mn. In (7), because each column of the matrix $\mathscr{A}_d$ is each column of A multiplied by nonzero scalar, $\mathrm{rank}(\mathscr{A}_d) = \mathrm{rank}(A)$. Number of free variables of the equation (6), which is degree of the null-space of $\mathscr{A}$, is $mn-(m-1)\mathrm{rank}(A) = (m-1)(n-\mathrm{rank}(A))+n$. Thus we obtain Theorem 1.

When $\mathrm{rank}(A) = n$ The equation $\mathscr{A}\mathbf{w} = 0$ is solved with Gaussian elimination which decomposes $A = LU$; the equation become $\mathscr{U}\mathbf{w} = 0$ which

$$\mathscr{U} = \begin{bmatrix} \mathscr{U}_1 & -\mathscr{U}_2 & & & \\ & \mathscr{U}_2 & -\mathscr{U}_3 & & 0 \\ & & \ddots & & \\ & 0 & & \mathscr{U}_{m-2} & -\mathscr{U}_{m-1} \\ & & & & \mathscr{U}_{m-1} \quad -\mathscr{U}_m \end{bmatrix}, \mathscr{U}_d = UA_d^{-1} = \left[\frac{1}{a_{d1}}\mathbf{u}_1 \ \cdots \ \frac{1}{a_{dn}}\mathbf{u}_n\right] \quad (8)$$

If $\mathrm{rank}(A) = n$ then there are n free variables. We can select n elements from $\mathbf{w}$ as free variables and can substitute any values for them. Let $\mathbf{B}^m$, solutions of last m equations in (6), be free variables; we substitute $\mathbf{t}' = [t'_1 \cdots t'_n]^\top \in \mathbb{R}^n_+$ for $\mathbf{B}^m$.

$$\mathbf{B}^m = \mathbf{t}', \quad \text{each element is } B^m_1 = t'_1, \quad B^m_2 = t'_2, \quad \cdots, \quad B^m_n = t'_n. \quad (9)$$

Other solutions are obtained by back-substitution.

$$UA_{m-1}^{-1}\mathbf{B}^{m-1} - UA_m^{-1}\mathbf{t}' = 0, \qquad \mathbf{B}^{m-1} = A_{m-1}A_m^{-1}\mathbf{t}'. \quad (10)$$

($\because$ when $\mathrm{rank}(U) = n$, $U(A_{m-1}^{-1}\mathbf{B}^{m-1} - A_m^{-1}\mathbf{t}') = 0$ iff $(A_{m-1}^{-1}\mathbf{B}^{m-1} - A_m^{-1}\mathbf{t}') = 0$.) $\mathbf{B}^i = A_iA_{i+1}^{-1}\mathbf{B}^{i+1}$ and $\mathbf{B}^{m-1}$ is in (10), thus more back-substitutions give

$$\mathbf{B}^i = A_iA_m^{-1}\mathbf{t}', \quad \text{or } B^i_j = \frac{a_{ij}}{a_{mj}}t'_j, \quad i = 1, \cdots, m-2. \quad (11)$$

Replace free variables as $\mathbf{t} = A_m^{-1}\mathbf{t}'$, or $[t_1, \cdots, t_n]^\top = \left[\frac{t'_1}{a_{m1}}, \cdots, \frac{t'_n}{a_{mn}}\right]^\top$,

$$\mathbf{B}^i = A_i\mathbf{t}, \quad \text{or } B^i_j = a_{ij}t_j, \quad i = 1, \cdots, m. \quad (12)$$

Finally, we obtain interchangeable criteria weight vectors by normalizations,

$$\mathbf{b}^i = \frac{\mathbf{B}^i}{\mathbf{1}^\top\mathbf{B}^i} = \frac{A_i\mathbf{t}}{\mathbf{1}^\top A_i\mathbf{t}}, \quad \text{or } b^i_j = \frac{a_{ij}t_j}{\sum_{\zeta=1}^n a_{i\zeta}t_\zeta}. \quad (13)$$

Theorem 2. *If* $\mathrm{rank}(A) = n$, *interchangeable criteria weight vectors* $\mathbf{b}^i$ *be expressed as* $\mathbf{b}^i = \frac{A_i\mathbf{t}}{\mathbf{1}^\top A_i\mathbf{t}}$. $\mathbf{t}$ *is any vector in* $\mathbb{R}^n_+$.

Then normalized aggregate scores are

$$\mathbf{e}^d = \frac{AA_d^{-1}\mathbf{b}^d}{\mathbf{1}^\top AA_d^{-1}\mathbf{b}^d} = \frac{AA_d^{-1}\frac{A_d\mathbf{t}}{\mathbf{1}^\top A_d\mathbf{t}}}{\mathbf{1}^\top AA_d^{-1}\frac{A_d\mathbf{t}}{\mathbf{1}^\top A_d\mathbf{t}}} = \frac{AA_d^{-1}A_d\mathbf{t}}{\mathbf{1}^\top AA_d^{-1}A_d\mathbf{t}} = \frac{A\mathbf{t}}{\mathbf{1}^\top A\mathbf{t}}, \tag{14}$$

$$\text{each element is } e_i^d = \frac{\sum_{j=1}^n a_{ij}t_j}{\sum_{\xi=1}^m\sum_{j=1}^n a_{\xi j}t_j}. \tag{15}$$

Aggregate scores are independent of whichever alternative we choose as dominant.

3 Discussions

In this section, we reconsider results of previous researches with our theorems.

Interchangeability and law of CRU A next theorem is derived from Theorem 2.

Theorem 3. *If* $\mathrm{rank}(A) = n$,

$$\mathbf{b}^k = \frac{A_kA_i^{-1}\mathbf{b}^i}{\mathbf{1}^\top A_kA_i^{-1}\mathbf{b}^i}. \tag{16}$$

Proof. $\mathbf{b}^i = \frac{A_i\mathbf{t}}{\mathbf{1}^\top A_i\mathbf{t}}$ ($\because$ Theorem 2), $\frac{A_kA_i^{-1}\mathbf{b}^i}{\mathbf{1}^\top A_kA_i^{-1}\mathbf{b}^i} = \frac{A_kA_i^{-1}\frac{A_i\mathbf{t}}{\mathbf{1}^\top A_i\mathbf{t}}}{\mathbf{1}^\top A_kA_i^{-1}\frac{A_i\mathbf{t}}{\mathbf{1}^\top A_i\mathbf{t}}} = \frac{A_k\mathbf{t}}{\mathbf{1}^\top A_k\mathbf{t}} = \mathbf{b}^k$. □

This theorem extrapolates a criteria weight vectors $\mathbf{b}^k$ from $\mathbf{b}^i$. The equation (16) is introduced as a premise, referred to as *law of CRU (Constant of Ratio in Unit Evaluation)*, of Dominant AHP in [1][2].

Theorem 3 says that only interchangeability does not keep the law. From Theorem 1 and Theorem 2, if $\mathrm{rank}(A) < n$ then a pair of vectors expressed as $\mathbf{b}^{i'} = \frac{A_{i'}\mathbf{t}+\delta^{i'}}{\mathbf{1}^\top[A_{i'}\mathbf{t}+\delta^{i'}]}$ and $\mathbf{b}^k = \frac{A_k\mathbf{t}}{\mathbf{1}^\top A_k\mathbf{t}}$ may exist. The equation (16) cannot estimate $\mathbf{b}^k$ from $\mathbf{b}^{i'}$.

$$\frac{A_kA_{i'}^{-1}\mathbf{b}^{i'}}{\mathbf{1}^\top A_kA_{i'}^{-1}\mathbf{b}^{i'}} = \frac{A_kA_{i'}^{-1}\frac{A_{i'}\mathbf{t}+\delta^{i'}}{\mathbf{1}^\top[A_{i'}\mathbf{t}+\delta^{i'}]}}{\mathbf{1}^\top A_kA_{i'}^{-1}\frac{A_{i'}\mathbf{t}+\delta^{i'}}{\mathbf{1}^\top[A_{i'}\mathbf{t}+\delta^{i'}]}} = \frac{A_k\mathbf{t}+A_kA_{i'}^{-1}\delta^{i'}}{\mathbf{1}^\top[A_k\mathbf{t}+A_kA_{i'}^{-1}\delta^{i'}]} \neq \mathbf{b}^k. \tag{17}$$

When, of course, we adjust $\delta^{i'}$ to 0 then $\mathbf{b}^{i'} = \mathbf{b}^k$.

Corollary 1. *A sufficient condition for holding the law of CRU is that criteria weight vectors are interchangeable and* $\mathrm{rank}(A) = n$.

Solutions of CCM CCM(Concurrent Convergence Method)[3] searches for interchangeable criteria weight vectors. A decision maker inputs $\beta^1, \cdots, \beta^m$, not interchangeable, into CCM as initial vectors. Algorithm 1 describes procedures of CCM.

CCM must output interchangeable vectors[4], and then from Theorem 2,

Algorithm 1 CCM

Input: $\{\beta^1,\cdots,\beta^m\}$, ε
Output: $\{\hat{\mathbf{b}}^1_{CCM},\cdots,\hat{\mathbf{b}}^m_{CCM}\}$.
$\mathbf{b}^k_0 := \beta^k$, $k = 1,...,m$.
$t := 0$.
repeat
$\mathbf{b}^i_{t+1} := \frac{1}{m}\sum_{k=1}^m \frac{A_i A_k^{-1}\mathbf{b}^k_t}{\mathbf{1}^\top A_i A_k^{-1}\mathbf{b}^k_t}$. (This state averages vectors.)
until $\max_{k\in(1,...,m)}||\mathbf{b}^k_{t+1}-\mathbf{b}^k_t|| < \varepsilon$
$\hat{\mathbf{b}}^k_{CCM} := \mathbf{b}^k_t$, $k = 1,...,m$.

Theorem 4. *If* $\mathrm{rank}(A) = n$, *a n-dimensional vector* $\mathbf{t}_{CCM}$ *exists and CCM outputs criteria weight vectors expressed as:*

$$\hat{\mathbf{b}}^i_{CCM} = \frac{A_i\mathbf{t}_{CCM}}{\mathbf{1}^\top A_i\mathbf{t}_{CCM}}, \quad i = 1,\cdots,m. \tag{18}$$

Interchangeable criteria weight vectors exist infinitely(Theorem 1), hence searching for these vectors is a multidimensional optimization problem. Because the set of vectors $\hat{\mathbf{b}}^1_{CCM},\cdots,\hat{\mathbf{b}}^m_{CCM}$ is one of infinite solutions, we need to check whether the vectors is the optimum or is not. We introduce two simple measurements,

$$dist(\hat{\mathbf{b}}^1,\cdots,\hat{\mathbf{b}}^m;\beta^1,\cdots,\beta^m) = \sum_{j=1}^n\sum_{i=1}^m(\hat{b}^i_j-\beta^i_j)^2, \tag{19}$$

$$dist^*(\hat{\mathbf{b}}^1,\cdots,\hat{\mathbf{b}}^m;\beta^1,\cdots,\beta^m) = \sum_{j=1}^n\sum_{i=1}^m\sum_{k=1}^m\left(\frac{\frac{a_{ij}}{a_{kj}}\hat{b}^k_j}{\mathbf{1}^\top A_i A_k^{-1}\hat{\mathbf{b}}^k} - \frac{\frac{a_{ij}}{a_{kj}}\beta^k_j}{\mathbf{1}^\top A_i A_k^{-1}\beta^k}\right)^2 \tag{20}$$

These are square of distance between outputs $\hat{\mathbf{b}}^i$ and inputs β^i. $dist^*$ has information of distance among aggregate scores with every vectors estimated by the law of CRU.

There is a numerical example that a CCM's output is not the optimum solution. [1]

$$\text{Inputs are } A = \begin{bmatrix} 1 & 1 \\ \frac{3}{8} & 32 \\ \frac{1}{8} & \frac{1}{16} \end{bmatrix}, \beta^1 = \begin{bmatrix}0.520\\0.480\end{bmatrix}, \beta^2 = \begin{bmatrix}0.510\\0.490\end{bmatrix}, \beta^3 = \begin{bmatrix}0.530\\0.470\end{bmatrix}, \tag{21}$$

$$\text{outputs are } \hat{\mathbf{b}}^1_{CCM} = \begin{bmatrix}0.775\\0.225\end{bmatrix}, \hat{\mathbf{b}}^2_{CCM} = \begin{bmatrix}0.039\\0.961\end{bmatrix}, \hat{\mathbf{b}}^3_{CCM} = \begin{bmatrix}0.865\\0.135\end{bmatrix}, \tag{22}$$

$$\text{and measurements are } dist = 0.798, dist^* = 1.315. \tag{23}$$

[1] These values of A, β^i, and $\hat{\mathbf{b}}^i_{CCM}$ are in Appendix 1 of [5]. $dist$ and $dist^*$ are calculated by us.

In this example, the optimum set of interchangeable vectors is [2]

$$\hat{\mathbf{b}}^1_{ref} = \begin{bmatrix} 0.615 \\ 0.385 \end{bmatrix}, \hat{\mathbf{b}}^2_{ref} = \begin{bmatrix} 0.018 \\ 0.982 \end{bmatrix}, \hat{\mathbf{b}}^3_{ref} = \begin{bmatrix} 0.762 \\ 0.238 \end{bmatrix}, \tag{24}$$

$$\text{and measurements are } dist = 0.609, dist^* = 1.151. \tag{25}$$

Normalizations in the averaging vectors in Algorithm 1 cause nonlinearity against variation of $\mathbf{b}^i_t$, and this makes outputs of CCM be not the optimum.

A relation between ANP and Dominant AHP Let us consider that evaluation matrix A and interchangeable criteria weight vectors $\mathbf{b}^1, \cdots, \mathbf{b}^m$ of Dominant AHP be inputs of ANP(Analytic Network Process)[6]. In this case, a solution of ANP is a pair of vectors $\mathbf{x}$ and $\mathbf{y}$ that hold next equation.

$$\begin{bmatrix} 0 & \mathbf{b}^1 \cdots \mathbf{b}^m \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \quad \mathbf{x} = [x_1, \cdots, x_n]^\top, \mathbf{y} = [y_1, \cdots, y_m]^\top. \tag{26}$$

Let $\text{rank}(A) = n$. Because $\mathbf{b}^i$ is expressed as $\frac{A_i \mathbf{t}}{\mathbf{1}^\top A_i \mathbf{t}}$ with a parameter $\mathbf{t} = [t_1, \cdots, t_n]^\top$ (Theorem 2), if we put as $x_j = t_j$ and $y_i = \sum_{j=1}^n a_{ij} t_j$ then the equation is

$$\begin{bmatrix} & & \frac{a_{11}t_1}{\sum_{j=1}^n a_{1j}t_j} & \frac{a_{21}t_1}{\sum_{j=1}^n a_{2j}t_j} & \cdots & \frac{a_{m1}t_1}{\sum_{j=1}^n a_{mj}t_j} \\ & 0 & \frac{a_{12}t_2}{\sum_{j=1}^n a_{1j}t_j} & \frac{a_{22}t_2}{\sum_{j=1}^n a_{2j}t_j} & \cdots & \frac{a_{m2}t_2}{\sum_{j=1}^n a_{mj}t_j} \\ & & \vdots & & \ddots & \vdots \\ & & \frac{a_{1n}t_n}{\sum_{j=1}^n a_{1j}t_j} & \cdots & & \frac{a_{mn}t_n}{\sum_{j=1}^n a_{mj}t_j} \\ a_{11} & a_{12} \cdots a_{1n} & & & & \\ a_{21} & a_{22} \cdots a_{2n} & & & & \\ \vdots & \ddots \vdots & & & 0 & \\ a_{m1} & \cdots a_{mn} & & & & \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \\ \sum_{j=1}^n a_{1j}t_j \\ \sum_{j=1}^n a_{2j}t_j \\ \vdots \\ \sum_{j=1}^n a_{mj}t_j \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m \left(\frac{a_{i1}}{\sum_{j=1}^n a_{ij}t_j} \sum_{j=1}^n a_{ij}t_j \right) t_1 \\ \sum_{i=1}^m \left(\frac{a_{i2}}{\sum_{j=1}^n a_{ij}t_j} \sum_{j=1}^n a_{ij}t_j \right) t_2 \\ \vdots \\ \sum_{i=1}^m \left(\frac{a_{in}}{\sum_{j=1}^n a_{ij}t_j} \sum_{j=1}^n a_{ij}t_j \right) t_n \\ \sum_{j=1}^n a_{1j}t_j \\ \sum_{j=1}^n a_{2j}t_j \\ \vdots \\ \sum_{j=1}^n a_{mj}t_j \end{bmatrix}$$

This equation is valid when $\sum_{i=1}^m a_{i1} = 1, \cdots, \sum_{i=1}^m a_{in} = 1$ (or each column of A is normalized), and then $\mathbf{y}$ in (26) corresponds to aggregate scores of Dominant AHP (When we normalize $\mathbf{y}$, each element of $\mathbf{y}$ equals e^d_i in (15)).

Theorem 5. *If the evaluation matrix A consists of normalized columns and* $\text{rank}(A) = n$, *aggregate scores and internal parameters of interchangeable weight vectors of Dominant AHP are solutions of ANP.*

This is a variety of the statement that is expressed symbolically in 'Dominant AHP + ANP = Dominant AHP' at [7].

[2] Each $\hat{\mathbf{b}}^i_{ref}$ is $A_i\mathbf{t}$, and $\mathbf{t} = [0.615, 0.385]^\top$ is obtained by 2-dimensional Newton's method that searches for t_j which holds $\frac{\partial dist^*}{\partial t_j} = 0$.

4 Conclusion

Kinoshita and Sekitani [4] manifest a existence of interchangeable criteria weight vectors by proofing that CCM's outputs must converge. In this paper, we show that these vectors exist infinitely and they are expressed with $(m-1)(n-\mathrm{rank}(A))+n$ parameters. This result be an important tool for analyzing Dominant AHP.

We search for criteria weight vectors which hold only interchangeability. These vectors may violate the law of CRU. But some adjustments of parameters can enforce the law. If $\mathrm{rank}(A)=n$, even these adjustments are unnecessary(Corollary 1).

When $\mathrm{rank}(A)=n$, interchangeable criteria weight vectors have simple form. And if $n>m$ then $\mathrm{rank}(A)\neq n$ and the degree of freedom is larger than n. This means that we must notice not only redundancy among criteria but also difference between number of criteria and number of alternatives, and must calculate $\mathrm{rank}(A)$.

Whole process of proofs in $\mathrm{rank}(A)=n$ case, if each A_i holds $\mathrm{rank}(AA_i^{-1})=\mathrm{rank}(A)$, same results will be derived. This enables us to extend evaluation processes, which construct A, of Dominant AHP. The matrix A_i is diagonal because a element of a row of A represents a ratio to a corresponding element of a dominant row; they are valued by comparing each element. The row of A, however, can be valued as a vector transformation of the dominant row when the transformation keeps $\mathrm{rank}(A)$; the transformation matrix be A_i and be not diagonal. This be useful, for examples, when a decision maker indicates relation among alternatives as positional relation in a plane diagram.

In the equation (14), if $\mathbf{t}$ is normalized, aggregate scores of Dominant AHP equal the scores of AHP[8] which has $\mathbf{t}$ as criteria weight vector. On the other hand, if $\mathrm{rank}(A)=n$ and $\mathbf{b}$ is a criteria weight vector at AHP, then each $\mathbf{b}^i=\frac{A_i\mathbf{b}}{\mathbf{1}^\top A_i\mathbf{b}}$ be the interchangeable criteria weight vector of Dominant AHP with same aggregate scores of AHP. This causes similarities of vector convergence properties between some AHP's methods and Dominant AHP's such as in Theorem 5.

References

1. Kinoshita, E., Nakanishi, M.: A Proposal of a New Viewpoint in Analytic Hierarchy Process. Journal of Infrastructure Planning and Management IV-36 569, 1–8 (1997) (in Japanese)
2. Kinoshita, E., Nakanishi, M.: Proposal of New AHP Model in light of Dominant Relationship among Alternatives. Journal of the Operations Research Society of Japan 42(2), 180–197 (1999)
3. Kinoshita, E., Nakanishi, M.: A Proposal of "Concurrent Convergence" as a Processing Technique for Additional Data in the Dominant Alternatives Method. Journal of Infrastructure Planning and Management IV-42 611, 13–19 (1999) (in Japanese)
4. Kinoshita, E., Sekitani, K., Shi, J.: Mathematical Properties of Dominant AHP and Concurrent Convergence Method. Journal of the Operations Research Society of Japan 45(2), 198–213 (2002)
5. Sekitani, K., Ueta, H.: Remarks on the Concurrent Convergence Method for a Typical Mutual Evaluation System. Journal of the Operations Research Society of Japan 47(2), 82–95 (2004)

6. Saaty, T.L.: The Analytic Network Process. Expert Choice (1996)
7. Kinoshita, E., Sugiura, S.: The Relationship between Dominant AHP/CCM and ANP. In: Watada, J., Phillips-Wren, G., Jain, L.C., Howlett, R.J., et al. (eds.) Intelligent Decision Technologies. SIST, vol. 10, pp. 319–328. Springer, Heidelberg (2011)
8. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

Decision Making on the Optimization of Budget Allocation for Safety

Yuji Sato

Abstract. Based on the quantitative model of risk assessment quantifying intangible factors in risk assessment, optimizing a budget allocation for safety in a manufacturing company was the objective of this paper.Developing a quantitative model for risk assessment in conjunction with decision-making on strategic investments for safety is complicated because of the subjective factors that enter into the inspection of a plant within a company and the choice of appropriate safety measures for the plant. The author's previous study addressed a part of this problem, and showed how to quantify inherent risks within a plant.Based on the study, the issue of the optimization of the budget allocation for safety measures was addressed by approximating a budget allocation model as a linear programming problem in this paper. A case study was carried out in a chemical company in Japan as an example, which guided how to optimize budget allocation for safety in the company.

Keywords: budget allocation, degree of risk reduction, percentage complete, LP.

1 Introduction

With growing interest in global environmental issues, manufacturing companies need to take responsibility for reducing plant-based risks such as fires, explosions or leakages, given their potentially devastating human and environmental consequences. In addition to taking responsibility, companies need to adopt strategic investment in safety, which is inseparable from profit generation due to the increasing focus on accountability to stakeholders. For manufacturing companies, however, developing a budget allocation model is complicated, where the introduction of safety measures is critical to the success of risk management.

Yuji Sato
GraduateSchool of Policy Science, MieChukyoUniversity, 1846,
Kubo, Matsusaka, Mie, 515-8511 Japan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 447–455.
springerlink.com

In designing such an allocation model for a plant, risk assessment by safety supervisors should form the core. Taking into account intangible factors such as sound from impact test, as well as objective plant data such as temperature of tanks, are essential. In addition, safety managers must evaluate and choose safety measures in making decisions on strategic investments which are usually costly and surrounded by uncertainty. Managers in manufacturing companies continually face such difficulties.

Sato (2011) focused on risk evaluation for a manufacturing company, and the correlation between the introduction of safety measures and the degree of risk reduction were clarified by applying the Analytic Hierarchy Process (AHP). Inherent risks of a plant within the company are quantified based on both concrete measures for risk reduction and the consensus of safety supervisors in the plant. In addition, the degree of risk reduction is evaluated based on the importance and the percentage complete of safety measures for risk reduction.

Another issue remains; the evaluation and choice of appropriate safety measures, which is often beset by uncertainty over the likely effects of the different measures available, particularly given the rapidly changing technological environment. Thus, the design of a budget allocation model and decision-making for budget allocation, relying heavily as it does on experience, knowledge, as well as intuition, means that the evaluation and choice of appropriate safety measures often lacks transparency and traceability.

The number of studies on budget allocation for safety measures within manufacturing companies is limited because of inherent plant-based risks, which results in a lack of a “standard” scheme of optimization for the budget allocation. Manufacturing companiessurrounded by uncertainties such as unexpected loss of disasters and safety measures’ envisagedeconomic effects, therefore, have tried to develop “haute couture” budget optimization by, for instance, consulting professional analysts for safety. While from the perspective of the uncertainty of the results of the investments, such as potential benefits or unexpected costs of early stage manufacturing technology, a budget allocation for safety measures is similar to that for new technologies in manufacturing, which serves as a useful reference for this issue. As Beaumont(1999) ascertained, the criteria that firms use to make investment decisions in manufacturing technology are: how firms manage the introduction of new technology; whether firms experience unanticipated effects from new technology; what factors impede or assist its implementation. Some of these factors cannot be clarified before the investmentsare implemented.

Although decision makers on investments are not completely ignorant of what the future might be, investment decisions are made under conditions of uncertainty. Frank (1998) considered the nature and acceptable level of risk, together with management’s personal attitude to risk. O’Brien and Smith (1993) also noted that investments in advanced manufacturing systems must be made while taking into account factors that are difficult to predict. In their papers, how a decision process was designed and managed was discussed, and theapplication of the AHP was proposed. Most paststudies, including Frank (1998) and O’Brien and Smith (1993), however, focused solely on the efficacy and efficiency of each safety measure in conjunction with cost minimization; sustainability of a company was not taken into account.

This paper aims to address thisissue;optimization of a budget allocation for safety, making allowances for accountability to stakeholders. Unlike paststudies, the optimization process of budget allocation proposed in this paper considers not only cost minimization but also the sustainability of a company, and offers quantitative measurement to the budget plan for safety measures. The optimization of a budget allocation is accomplished by a traditional technique for optimization, Linear Programming (LP), which is used to optimize the budget allocation for safety measures in terms of the criteria of both sustainability and cost minimizationof a company. As an example of manufacturing companies, a case in a Japanese chemical company (whose name cannot be disclosed due to confidentialityagreement) is the basis of this paper.

2 Quantitative Model for Risk Assessment

This section briefly introducesa quantitative model for risk assessment. The model was proposed by Sato (2011) which succeeded in quantifying intangible factors in risk assessment. The quantification of risks within the chemical plant is undertaken using the following two steps, combining the breadth and depth of risks within the plant and incorporating quantitative and qualitative judgments in risk assessment: (*i*) identify and classify risks according to past experience; (*ii*) develop a quantitative model of risk assessment reflecting safety supervisors' perceptions. Omitting the details of the procedure here, they conducted pairwise comparisons over all possible combinations of dimensions and measures for risk reduction, such as "Which dimension do you think is more important for the risk reduction of our plant, Equipment or Regulation?"

Let ={} and ={}, respectively denote the weight of dimensions and that of measures of a risk reduction plan elicited by safety supervisors by using the AHP, where *D* and *M* show a set of dimensions and measures, respectively. Further let *P*={} denote percentage complete of each measure, which represents the achievement record of each safety measure in comparison with the initial estimate each year. Then the degree of risk reduction can be defined as follows;

$$\text{Degree of risk reduction:}= . \tag{1}$$

3 Optimization of Budget Allocation

In this section, the optimization scheme of budget allocation is proposed based on the quantitative model of risk assessment introduced in the previous section. To optimize a budget allocation for safety, chemical companies must consider not only necessary and sufficient expenditures but also their social responsibility so as to avoid accidents that might seriously damage the environment. That is, optimizing a budget allocation for safety measures entails not only minimizing costs for risk management but also integrating the economic, legal and social engineering perspectives within the framework. In this paper, LP is applied to settle the issue of budget optimization. Linear functions describing the relationship between expenditures for safety measures and the percentage

complete of each safety measure is formulated. Then by solving the LP, the budget allocation for safety of a company is optimized.

The rationale for the linear approximation is the difficulty in identifying a function representing the relationship between the expenditures and the percentage complete. The relationship might be a non-linear function, and while even a non-linear programming problem could be solved mathematically, it is not easy to precisely identify a function as well as to solve the non-linear programming problem for practitioners. On the other hand, every function can be represented as a linear function by piecewise linear approximation, which implies the decomposition of risk reduction plans (processes) in the optimization of the budget allocation. In addition, dealing with safety measures based on many years of experience is likely to be more familiar for practitioners than precisely identifying the function with abstruse mathematics. Thus, the linear approximation of the relationship and the use of the function in optimizing budget allocation in a real case could be justified. These approaches would be easily understood by practitioners, thus enhancing the applicability of the scheme.

On the other hand, the LP has budget constraints as well as non-negative conditions, which imposes a ceiling on the expenditures for safety measures. In the event of accidents, companies would fully expense what they must cover for restoration to the original state in order to assume responsibility for the damage they had caused. The budget allocation "plan" for safety, however, includes preventive measures for safety as well as countermeasures based on the estimation as to how much they should expense for future reliability, which are not always necessary expenses at a particular budgeting. Furthermore, companies need to invest not only for safety but also for items such as R&D and welfare for employees; they must "balance" their budget allocation for each field they are responsible for. Thus, optimization of the budget allocation would have upper-limit constraints.

The optimization of the budget allocation for safety is carried out using the following three steps, evaluating the efficacy of safety measures and formulating LP which optimizes budget allocation for the safety measures.

Step 1: An evaluation of the efficacy of safety measures for the chemical plant is carried out based on the quantitative model of risk assessment developed in the previous study (Sato, 2011). The evaluation clarifies the correlation between the impact of each safety measure and the degree of risk reduction. The AHP will also be applied to this process since pairwise comparison functions better than traditional questionnaire methods in quantifying human perception (Sato, 2004, 2009).

Step 2: The relationship between expenditures and the degree of risk reduction of each safety measure is linearly approximated. Given resources such as human, material and budget for each safety measure, safety supervisors could estimate percentage complete of each measure.

Step 3: Based on the degree of risk reduction calculated from the model, an LP problem optimizing a budget allocation for safety is formulated. The LPconsists

of an objective function maximizing the degree of risk reduction with given resources (*i.e.*, cost minimizing of budget amounts for safety measures with allowable risk level) and of constraints which include not only budget constraints but also the constraint concerning the sustainability of a company. The solution is derived from the LP when the actual data would be collected from the chemical plant.

Let denote the expenditures for a measure Suppose the percentage complete can be approximated by a linear function of the budget amount for each safety measure $L()$, then "maximizing the degree of risk reduction" problem with budgetconstraint for each measure would be formulated as the LP problem as shown below.

$$\text{Max } \{\}, \tag{2}$$

$$\text{Subject to} \leq , \tag{3}$$

$$0 \leq , \tag{4}$$

wheredenotes the budget constraint of dimension .

The objective function (2) of this LP consists of the degree of risk reduction, which is the product of the weight of each dimension, that of each measure and the linear approximation of the percentage complete of each measure $L()$. Both and are derived from the quantitative model developed in the previous study, and $L()$ is estimated by the safety supervisors based on given resources. The linear constraints (3) and (4) are the budget constraint of each dimension and non-negative condition. This LP has, by nature, a convex feasible region, and thus derives optimal solutions, which in turn optimizes the budget plan for safety measure for the chemical plant.

4 A Case Study in a Chemical Company in Japan (Example)

This section demonstrates an application of the budget allocation scheme proposed in the previous section which is based on the quantitative model of risk assessment used in a Japanese chemical company. Although the chemical company has tried to reduce the number of accidents, the current safety status of the plant is not satisfactory. The number of accidents increased 80% from 2007 to 2008, after they succeeded in reducing the number of accidents in 2007. To reduce risks, the company undertook case studies for shareholders, and consulted a professional analyst for the safety of chemical plant. Based on the classification, concrete measures for risk reduction were proposed as shown in Table 1, consisting of company-wide projects and activities within the plant. The quantitative model of risk assessment in conjunction with the risk reduction scheme was developed based on the safety measures listed in Table 1.

The numbers in Table 1 represent the weights for dimensions and measures normalized by the -norm within each cluster which are obtained from the previous study. As shown in the table, the safety supervisors weighted "Human" most, 0.390, among three dimensions. For the measures, "Tank preservation" ranked

highest, 0.201, among the measures for Equipment; “OHSMS” (Occupational Health and Safety Management System) ranked highest, 0.303, among those for Human; and “Fire code observance” ranked highest, 0.329, among those for Regulation.

The scheme of budget planning for safety measures is based on the abovementioned actual quantitative model of risk assessment, and the grades employed in the following part are examples. Based on Table 1, which summarizes the quantitative model of risk assessment for the chemical plant, the objective function maximizing the degree of risk reduction can be described as (2), where the percentage complete is approximated by $L()$, the linear functions of the budget amount for each measure. The linear constraints (3), (4) are the budget constraints for each dimension and non-negative condition, where i= 1, …, 3; j=1, …, 7.

Table 1 Safety measures for risk reduction.

Dimension ()	Safety measures ()	
	Projects	Activities
Equipment (0.365)	Emergency inspection (0.161) Static electricity measure (0.135) Tank preservation (0.201) FMEA (0.114)	Electricity intentional preservation (0.190) Environmental risk hedge activity (0.124) Incinerator abolition (0.0745)
Human (0.390)	Superintendent arousal (0.190) Security intensification (0.186)	Natural calamity measure (0.161) License institution (0.161) OHSMS (0.303)
Regulation (0.245)	Equipment measurement system (0.196) Fire code observance (0.392) Inspection system establishment (0.128)	Zero-emission activity (0.130) 5S activity (0.217)

In this paper, the parameters defining the linear functions $L()$ and the budget constraints of each dimension are assumed to be those listed in Table 2. The constant in the table denotes necessary and sufficient expenditures to achieve 100% complete for each safety measure estimated by the safety supervisors. Since some measures for risk reduction are related to one another, the secondary effect of the expenditures for the primary measure to the percentage complete of the related measure is assumed, for this paper, to be half of the ratio 1/of the primary measure. Then the parameters of the LP optimizing budget plan of the company can be represented as shown in Table 2.

Table 3 summarizes the optimal solution derived from the LP. The numbers in the andcolumns respectively represent the optimal expenditures for each measure and its estimated percentage complete. In terms of the risk level, the degree of risk reduction under the obtained budget plan is improved from 0.392 (cf., Sato, 2011) to 0.588, which implies the estimated risk level one year later would be reduced from 0.608 (cf., Sato, 2011) to 0.412. This result just shows an example of the optimization of a budget plan. The process, on the other hand, is based on the actual quantitative model of risk assessment in a Japanese chemical company, which could easily be applied to an actual budget allocation after the linear approximation of the percentage complete of each measure.

Table 2 Framework of the LP maximizing the degree of risk reduction.

Dimension	Measures				1/	related measures	= *L*()
Equipment =0.365 =250 #	Emergency inspection	()	0.161	150	0.00667		0.00667 + 0.005
	Static electricity measure	()	0.135	75	0.0133	-	0.0133
	Tank preservation	()	0.201	80	0.0125	-	0.0125
	FMEA	()	0.114	100	0.0100	-	0.01
	Electricity intentional preservation	()	0.190	110	0.00909	,	0.00665 +0.005 + 0.00909
	Environmental risk hedge activity	()	0.124	50	0.0200	,	0.00333 +0.005 + 0.02
	Incinerator abolition	()	0.0745	150	0.00667	,	0.00625 +0.005 + 0.0667
Human =0.390 =50 #	Superintendent arousal	()	0.190	50	0.0200	-	0.02
	Security intensification	()	0.186	75	0.0133	-	0.0133
	Natural calamity measure	()	0.161	50	0.0200	,	0.01 +0.00665 + 0.02
	License institution	()	0.161	25	0.0400	,	0.01 +0.00665 + 0.04
	OHSMS	()	0.303	50	0.0200	,	0.01 +0.00665 + 0.02
Regulation =0.245 =35 #	Equipment management system	()	0.196	15	0.0667	-	0.0667
	Fire code observance	()	0.329	10	0.100		0.005 + 0.1
	Inspection system establishment	()	0.128	18	0.0556	-	0.0556
	Zero emission activity	()	0.130	14	0.0714	,	0.0333 +0.0278 + 0.0714
	"5S" activity	()	0.217	13	0.0769	,	0.0333 +0.0278 + 0.0769

#: million JPY

Table 3 Results of the optimization.

Dimension	Measures			**
Equipment =250 #	Emergency inspection	30.4	0.355	0.0209
	Static electricity measure	30.1	0.402	0.0198
	Tank preservation	69.3	0.867	0.0635
	FMEA	15.3	0.153	0.00638
	Electricity intentional preservation	49.6	1.000	0.0694
	Environmental risk hedge activity	50.0	1.000	0.0452
	Incinerator abolition	5.29	0.0947	0.00258
			subtotal:	**0.228**
Human =50 #	Superintendent arousal	0.000	0.000	0.000
	Security intensification	0.000	0.000	0.000
	Natural calamity measure	5.54	0.203	0.0127
	License institution	16.5	0.937	0.0587
	OHSMS	27.9	1.000	0.118
			subtotal:	**0.190**
Regulation =35 #	Equipment management system	1.66	0.111	0.00532
	Fire code observance	10.0	1.000	0.0806
	Inspection system establishment	0.000	0.000	0.000
	Zero emission activity	10.3	1.000	0.0318
	"5S" activity	13.0	1.000	0.0532
			subtotal:	**0.171**
#: million JPY			**Degree of risk reduction:**	**0.588**
			Estimated risk level 1 year later:	**0.412**

5 Concluding Remarks

This paper focuses on an optimization of a budget allocation for safety in a manufacturing company based on the quantitative model of risk evaluation of the plant in the company.The correlation between safety measures and the degree of risk reduction are clarified based on both concrete measures for risk reduction and the consensus of safety supervisors in the plant. In addition, the degree of risk reduction is evaluated based on the importance and the percentage complete of safety measures for risk reduction. Furthermore, not only the cost effectiveness of budget allocation for safety but also the effects of investments for risk reduction on the sustainability of a company are considered, thus optimizing the budget plan for safety measures.

The followingsareopen-ended questions. First, how to approximate the percentage complete of each safety measure for actual cases within a chemical company?—the relationship between budget amounts and percentage complete of each safety measure just might not lend itself to linear approximation. In such

cases, the process of a safety measure needs to be segmented so as to formulate an LP problem. Second, how to promote horizontal development in other cases?—the results obtained from this paper is just a case, which is valid only in the company surveyed. More robust model which can be applied to other companies must be explored.

References

Beaumont, N.B.: Investment decisions in Australian manufacturing. Technovation 18(11), 689–695 (1999)

Frank, L.: Approaches to risk and uncertainty in the appraisal of new technology capital projects. International Journal of Production Economics 53(1), 21–33 (1998)

O'Brien, C., Smith, S.J.E.: Design of the decision process for strategic investment in advanced manufacturing systems. International Journal of Production Economics 30-31, 309–322 (1993)

Sato, Y.: Comparison between Multiple-choice and Analytic Hierarchy Process: Measuring Human Perception. International Transactions in Operational Research 11(1), 77–86 (2004)

Sato, Y.: How to Measure Human Perception in Survey Questionnaires. International Journal of the Analytic Hierarchy Process 1(2), 64–82 (2009)

Sato, Y.: A Quantitative Model for Budget Allocation for Investment in Safety Measures. In: Watada, J., Phillips-Wren, G., Jain, L.C., Howlett, R.J. (eds.) Intelligent Decision Technologies. SIST, vol. 10, pp. 57–64. Springer, Heidelberg (2011)

Derivations of Factors Influencing the Technology Acceptance of Smart TV by Using the DEMATEL Based Network Process

Chi-Yo Huang, Yu-Wei Liu, and Gwo-Hshiung Tzeng

Abstract. Information technology (IT) products have played significant roles from both aspects of market development as well as technology growth. However, predicting technology acceptance toward the IT products is always challenging for marketers and designers. Also, the characteristics of and strategies for IT products at different stages of product life cycle are different. Therefore, the customers' preference and needs change rapidly at different stages in product life cycle. However, very few scholars tried to explore how the technology can be accepted at various stages of a product life cycle was seldom discussed. But the issue is very important. Thus, this research aims to propose DEMATEL based Network Process for discovering the influential factors of the technology acceptance for introduction stage IT product. An empirical study based on the Smart TV for introduction stage will be used to verify the feasibility of this framework. Further, the Lead User Method will be applied to predict the preference and the purchase behavior of mass user and the result could be a reference for industrial marketing activities and design process.

Keywords: Technology Acceptance Model (TAM), Lead User Method, DEMATEL based Network Process (DNP), Multiple Criteria Decision Making (MCDM).

Chi-Yo Huang · Yu-Wei Liu
Department of Industrial Education, National Taiwan Normal University
No. 162, Hoping East Road I, Taipei 106, Taiwan
e-mail: cyhuang66@ntnu.edu.tw

Gwo-Hshiung Tzeng
Graduate Institute of Project Management, Kainan University, No. 1, Kainan Road, Luchu, Taoyuan County 338, Taiwan
and
Institute of Management of Technology, National Chiao Tung University, Ta-Hsuch Road, Hsinchu 300, Taiwan
e-mail: ghtzeng@mail.knu.edu.tw, ghtzeng@cc.nctu.edu.tw

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 457–469.
springerlink.com

1 Introduction

During the past decades, social and personality psychologists attempted to study human behaviors. However, considering the complexity, explanation and prediction of human behavior is a difficult task. The concepts referring to behavioral intentions like the social attitude, personal intention and personality traits have played important roles in prediction of human behaviors (Ajzen 1989; Bagozzi 1981).

The Smart TV combined the simplicity of the normal TV and remote experience with the powerful search features and video on demand libraries you're accustomed to on the Web (Patrick, 2011). Toss in social networking, photo sharing, music, gaming and a hundred kinds of Web content (Patrick, 2011). Never needing to settle for anything less than having what you want to watch (or hear, or play) running in big-screen glory while you master the universe from the couch with the all-powerful remote (Patrick, 2011). Analyzing and predicting the consumer behaviors of Smart TVs for fulfilling customers' needs has become an indispensable task for marketing managers. However, the predictions are not easy. The consumer electronics technology evolved rapidly which shorten the product life cycles. How the consumers' acceptance of novel Smart TVs can be analyzed and predicted have become an important issue for marketing managers of Smart TV providers. However, the analysis and prediction of consumer behaviors are not easy from the above aspects mentioned aspects on PLC, competition, and alternative products.

In order to accurately derive the factors influencing consumers' acceptance of Smart TVs and predict the purchase behaviors, the Technology Acceptance Model (TAM) and the Lead User Method (LUM) will be introduced as the theoretic basis of this analysis. The TAM can be used for illustrating the factors influencing the consumers' acceptance of future Smart TVs. Meanwhile, the LUM is more suitable for evaluating the future high technology innovations which are disruptive in nature.

Consequently, a novel DEMATEL based network process (DNP) based multiple criteria decision making (MCDM) framework will be proposed for deriving the acceptance intention of Smart TVs. The criteria for evaluating factors influencing the acceptance of the Smart TVs will first be summarized by the literature review. Then, the casual structure corresponding to the acceptance intention of Smart TVs will be derived by using the DNP methods. The structure versus each criterion from lead users' perspective will be established by using the DNP. Then, the influential criteria weights of lead users will be calculated by the DNP method.

An empirical study on the feasibility of the DNP based LUM on Smart TVs' acceptance predictions will be derived by the opinions of three lead users now serving in the world's leading IT firms. The results can serve as the basis for marketers' predicting of consumers' intentions of accepting the Smart TVs. Meanwhile, such knowledge can serve as the foundation for future marketing strategy definitions.

The remainder of this article is organized as follows. The related literature regarding to technology acceptance theories and TAM model will be reviewed in Section 2. The analytic framework based on the DNP framework will be introduced in Section 3. Then, in Section 4, an empirical study follows, designing the Smart TV on the proposed DNP based TAM framework. Managerial implications as well as discussions will be presented in Section 5. Finally, the whole article will be concluded in Section 6.

The Smart TV combined the simplicity of the normal TV and remote experience with the powerful search features and video on demand libraries you're accustomed to on the Web (Patrick, 2011). Toss in social networking, photo sharing, music, gaming and a hundred kinds of Web content (Patrick, 2011). Never needing to settle for anything less than having what you want to watch (or hear, or play) running in big-screen glory while you master the universe from the couch with the all-powerful remote (Patrick, 2011). Analyzing and predicting the consumer behaviors of Smart TVs for fulfilling customers' needs has become an indispensable task for marketing managers. However, the predictions are not easy. The consumer electronics technology evolved rapidly which shorten the product life cycles. How the consumers' acceptance of novel Smart TVs can be analyzed and predicted have become an important issue for marketing managers of Smart TV providers. However, the analysis and prediction of consumer behaviors are not easy from the above aspects mentioned aspects on PLC, competition, and alternative products.

In order to accurately derive the factors influencing consumers' acceptance of Smart TVs and predict the purchase behaviors, the Technology Acceptance Model (TAM) and the Lead User Method (LUM) will be introduced as the theoretic basis of this analysis. The TAM can be used for illustrating the factors influencing the consumers' acceptance of future Smart TVs. Meanwhile, the LUM is more suitable for evaluating the future high technology innovations which are disruptive in nature.

Consequently, a novel DEMATEL based network process (DNP) based multiple criteria decision making (MCDM) framework will be proposed for deriving the acceptance intention of Smart TVs. The criteria for evaluating factors influencing the acceptance of the Smart TVs will first be summarized by the literature review. Then, the casual structure corresponding to the acceptance intention of Smart TVs will be derived by using the DNP methods. The structure versus each criterion from lead users' perspective will be established by using the DNP. Then, the influential criteria weights of lead users will be calculated by the DNP method.

An empirical study on the feasibility of the DNP based LUM on Smart TVs' acceptance predictions will be derived by the opinions of three lead users now serving in the world's leading IT firms. The results can serve as the basis for marketers' predicting of consumers' intentions of accepting the Smart TVs. Meanwhile, such knowledge can serve as the foundation for future marketing strategy definitions.

The remainder of this article is organized as follows. The related literature regarding to technology acceptance theories and TAM model will be reviewed in Section 2. The analytic framework based on the DNP framework will be introduced in Section 3. Then, in Section 4, an empirical study follows, designing

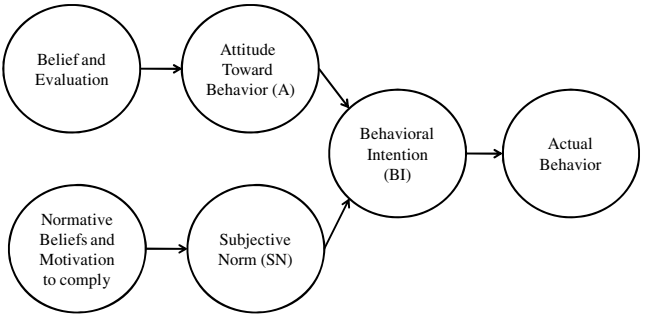
the Smart TV on the proposed DNP based TAM framework. Managerial implications as well as discussions will be presented in Section 5. Finally, the whole article will be concluded in Section 6.

2 Literature Review

The well-known consumer behavior analysis models such as TRA, TAM, Lead User Method and Product Life Cycle Theory which have been applied in many fields will be brief reviewed. Also, the product life cycle theory will be introduced.

2.1 Theory of Reasoned Action (TRA)

The theory of reasoned action (TRA) (Ajzen & Fishbein, 1980; Fishbein, 1980; Fishbein & Ajzen, 1975) has dominated research on attitude–behavior relations since its inception (Olson & Zanna, 1993). TRA is a prediction model refers to social psychology, which attempts to figure out the requirements of intended behaviors regarding to the acceptance of users. According to Fishbein and Ajzen (Davis, 1989; Hu & Bentler, 1999; Fishbein & Ajzen, 1974), user's specified behavior will be determined and influenced by users' behavioral intentions. Therefore, researchers applying the TRA should identify whether the belief is a prominent variable for users regarding the product acceptance behavior (Davis et al. 1989).



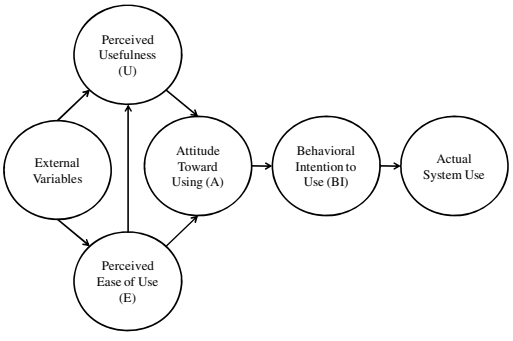
Source: Fishbein & Ajzen (1974)

Fig. 1 Theory of Reasoned Action (TRA)

2.2 Technology Acceptance Model (TAM)

The TAM, an adaptation of the TRA, was proposed especially for predicting the users' acceptance of information systems (Davis et al. 1989) by Davis (1986). The TAM is in general, capable of explaining user behaviors across a broad range of end-user computing technologies and user populations, while at the same time being both parsimonious and theoretically justified (Davis 1986; Davis et al. 1989). The research of TAM addresses the ability to predict peoples' computer

acceptance from a measure of their intentions, and the ability to explain their intentions in terms of their attitudes, subjective norms, perceived usefulness, perceived ease of use, and related variables (Davis, 1989; Davis, 1986).



Source: Davis (1986)

Fig. 2 Technology Acceptance Model

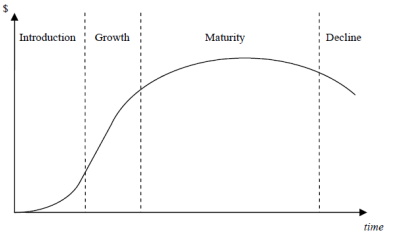
2.3 Lead User Method (LUM)

Lead-user theory is an approach which was originally proposed to selectively identify commercially attractive innovations developed by users (von Hippel, 1986). In previous paper, von Hippel (1986) has proposed that analysis of needs and solution data from "lead users" can improve the productivity of new product development in fields characterized by rapid change. Consequently, the LUM, a market analysis technique, is applied to the development of new products and services (Urban and Hippel 1988). The methodology is composed of four major steps based on the work of Urban and Hippel (1988): (1) specify lead user indicators, (2) identify lead user group, (3) generate concept (product) with lead users and (4) test lead user concept (product). Further details can be found in the earlier work by Urban and Hippel (1988).

2.4 Product Life Cycle Theory

The best-known and best-established theory of market maturation is the product life cycle theory pioneered by marketing scholars during the 1950s and 1960s (Yoo, 2010). Although variants of the product life cycle theory exist, the predominant version posits that the sales history of a new product will follow an S-shaped curve over time and that the product will pass through the four stages of sales growth depicted in Figure 3.

This research collected the criteria from the TAM for deriving the influential factors of technology acceptance of the Smart TV. To predict the mass users' preference of the Smart TV, this paper applied LUM and DNP method. The analytic framework will be introduced in the next section.



Source: Yoo (2010)

Fig. 3 The Product Life Cycle

3 Analytic Framework for Deriving the TAM

In order to build the analytical framework for comparing the factors influencing the acceptance of the Smart TVs from the aspects of lead users, the DNP based MCDM framework will be introduced. At first, the criteria being suitable for measuring the users' acceptance of Smart TVs will be derived based on literature review. The factors being identified by Davis in the TAM will be introduced. The DNP will then be introduced for deriving the causal relationship and influential weights versus each criterion from the lead users' aspect. Finally, the result will demonstrate the preference of lead users for predicting the mass customers' preference of Smart TV.

The DEMATEL based Network Process Method

The DNP is an MCDM framework consisting of the DEMATEL and the basic concept of ANP. The DEMATEL technique was developed by the Battelle Geneva Institute to analyze complex "real world problems" dealing mainly with interactive map-model techniques (Gabus and Fontela 1972) and to evaluate qualitative and factor-linked aspects of societal problems.

The DEMATEL technique was developed with the belief that the pioneering and proper use of scientific research methods could help to illuminate specific and intertwined phenomena and contribute to the recognition of practical solutions through a hierarchical structure. The ANP is general form of the analytic hierarchy process (AHP) (Saaty 1980, 1996) which has been used in MCDM problems by releasing the restriction of the hierarchical structure and the assumptions of independence between criteria. Combining the DEMATEL and basic concept of ANP method, which had been reviewed in this Section, the steps of this method can be summarized on the work of Prof. Tzeng, Gwo-Hshiung (Wei et al. 2010; Jeng & Tzeng, 2012):

Step 1: Calculate the direct-influence matrix by scores. Based on experts' opinions, the relationships between criteria can be derived based on mutual influences. The scale ranges from 0 to 4, representing "no influence" (0), "low influence" (1), "medium influence" (2), "high influence" (3), and "very high influence" (4), respectively. Respondents are requested to indicate the direct influence of a factor i on a factor j, or d_{ij}. The direct influence matrix $\boldsymbol{D}$ can thus be derived.

Step 2: Normalize the direct-influence matrix based on the direct-influence matrix D. The normalized direct relation matrix X can be derived by using as follow.

$$\boldsymbol{N} = v\boldsymbol{D};\ v = \min\left\{1/\max_i \sum_{j=1}^{n} d_{ij}, 1/\max_j \sum_{i=1}^{n} d_{ij}\right\},\ i, j \in \{1,2,...,n\}. \quad (1)$$

Step 3: Attain the total influence-related matrix T. Once the normalized direct-influence matrix N is obtained, the total influence-related matrix T can further be derived by using as following Eq. (2).

$$\boldsymbol{T} = \boldsymbol{N} + \boldsymbol{N}^2 + ... + \boldsymbol{N}^h = \boldsymbol{N}(\boldsymbol{I} - \boldsymbol{N})^{-1}, \text{ when } h \to \infty. \quad (2)$$

Then we can draw the network relation map (NRM) by using total influence-related matrix T; N is a direct influence matrix and $N = [x_{ij}]_{n\times n}$; $\lim_{h\to\infty}\left(\boldsymbol{N}^2 + \cdots + \boldsymbol{N}^h\right)$ stands for indirect influence matrix and $0 \le x_{ij} < 1$, and $0 < \sum_{j=1}^{n} x_{ij} \le 1$ and $0 < \sum_{i=1}^{n} x_{ij} \le 1$; If at least one row or column of summation is equal to 1, but not all, in $\sum_{j=1}^{n} x_{ij}$ or $\sum_{i=1}^{n} x_{ij}$ respectively; then we can guarantee $\lim_{h\to\infty} N^h = [0]_{n\times n}$. The element t_{ij} of the matrix T denotes the direct and indirect influences of the factor i on the factor j.

Step 4: Analyze the result. In this stage, the row and column sums are separately denoted as r and c within the total relation-related matrix T through Eq. (3).

$\boldsymbol{T} = [t_{ij}],\ i, j \in \{1,2,...,n\}$,

$$\boldsymbol{r} = [r_i]_{n\times 1} = \left[\sum_{j=1}^{n} t_{ij_n}\right]_{n\times 1} \text{ and } \boldsymbol{c} = [c_j]_{1\times n} = \boldsymbol{r}' = [r_i]'_{n\times 1} = \left[\sum_{i=1}^{n} t_{ij}\right]'_{1\times n} \quad (3)$$

where the r and c vectors denote the sums of the rows and columns, respectively.

Suppose the r_i denotes the row sum of the i^{th} row of the total relation-related matrix T. Then, r_i is the sum of the influences dispatching from the factor i to the all other factors, both directly and indirectly. Suppose that c_j denotes the column sum of the j^{th} column of the total relation-related matrix T. Then, c_j is the sum of the influences that factor i is receiving from the all other factors. Furthermore, when $i = j$ (i.e., the sum of the row sum and the column sum), $(r_i + c_i)$ represents the index representing the strength of the influence, both dispatching and receiving), where $(r_i + c_i)$ is the degree of the central role the factor i plays in the problem. If $(r_i - c_i)$ is positive, then the factor i primarily is dispatching influences upon the strength of other factors; and if $(r_i - c_i)$ is negative, then the factor i primarily is receiving influence from other factors. Therefore, a causal graph can be achieved by mapping the dataset of

$(r_i + s_i, r_i - s_i)$ providing a valuable approach for decision making to draw the map of influential relationship (Chiu et al. 2006; Huang and Tzeng 2007; Huang et al. 2011; Huang et al., 2007; Liou, Tzeng, & Chang, 2007; Yang & Tzeng, 2011; Shen et al., 2011; Ho et al., 2011; Liou & Tzeng, 2011; Chen & Tzeng, 2011; Tzeng & Huang, 2012; Liu et al., 2012; Wang & Tzeng, 2012; Ou Yang et al., 2012; Kuan et al., 2012).

Let the total influence-related matrix $\boldsymbol{T}_c = \left[\boldsymbol{T}_c^{ij}\right]_{n\times n}$ by criteria based on influence-related matrix $\boldsymbol{T}_c^{ij}$ of dimension *i* influence-related to dimension *j* as shown Eq. (4).

$$\boldsymbol{T}_c = \begin{array}{c} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_n \end{array} \begin{array}{c} \begin{array}{cccccc} c_{11} & \cdots c_{1m_1} & \cdots & c_{j1}\cdots c_{jm_j} & \cdots & c_{n1}\cdots c_{nm_n} \end{array} \\ \begin{bmatrix} \boldsymbol{T}_c^{11} & \cdots & \boldsymbol{T}_c^{1j} & \cdots & \boldsymbol{T}_c^{1n} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{i1} & \cdots & \boldsymbol{T}_c^{ij} & \cdots & \boldsymbol{T}_c^{in} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{n1} & \cdots & \boldsymbol{T}_c^{nj} & \cdots & \boldsymbol{T}_c^{nn} \end{bmatrix} \end{array} \begin{array}{l} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_n \end{array} \begin{array}{l} \begin{array}{l} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \end{array} \\ \begin{array}{l} c_{i1} \\ c_{i2} \\ \vdots \\ c_{im_i} \end{array} \\ \begin{array}{l} c_{n1} \\ c_{n2} \\ \vdots \\ c_{nm_n} \end{array} \end{array} \tag{4}$$

Step 4: Normalize total degree of influence-related matrix $\boldsymbol{T}_C$ by criteria in each dimension. Normalized matrix $\boldsymbol{T}_C^{\alpha}$ can be obtained from $\boldsymbol{T}_C$ by criteria in each dimension as Eq. (5).

$$\boldsymbol{T}_C^{\alpha} = \begin{array}{c} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_n \end{array} \begin{array}{c} \begin{array}{cccccc} c_{11}\cdots c_{1m_1} & \cdots & c_{j1}\cdots c_{jm_j} & \cdots & c_{n1}\cdots c_{nm_n} \end{array} \\ \begin{bmatrix} \boldsymbol{T}_c^{\alpha 11} & \cdots & \boldsymbol{T}_c^{\alpha 1j} & \cdots & \boldsymbol{T}_c^{\alpha 1n} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{\alpha i1} & \cdots & \boldsymbol{T}_c^{\alpha ij} & \cdots & \boldsymbol{T}_c^{\alpha in} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{\alpha n1} & \cdots & \boldsymbol{T}_c^{\alpha nj} & \cdots & \boldsymbol{T}_c^{\alpha nn} \end{bmatrix} \end{array} \begin{array}{l} \\ D_1 \\ \vdots \\ D_i \\ \vdots \\ D_n \end{array} \begin{array}{l} \begin{array}{l} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \end{array} \\ \begin{array}{l} c_{i1} \\ c_{i2} \\ \vdots \\ c_{im_i} \end{array} \\ \begin{array}{l} c_{n1} \\ c_{n2} \\ \vdots \\ c_{nm_n} \end{array} \end{array} \tag{5}$$

Step 5: Obtain un-weighted super-matrix in ANP. Un-weighted super-matrix $\boldsymbol{W}$ can be obtained from the transpose of normalized total degree of influence-related matrix $\boldsymbol{T}_C$ by criteria in each dimension using basic concept of ANP according to the interdependent relationship among dimensions and criteria.

$$\boldsymbol{W} = (\boldsymbol{T}_c^{\alpha})' = \begin{array}{c} \\ D_1 \\ \vdots \\ D_j \\ \vdots \\ D_n \end{array} \begin{array}{c} \begin{array}{cccccc} c_{11}\cdots c_{1m_1} & \cdots & c_{i1}\cdots c_{im_i} & \cdots & c_{n1}\cdots c_{nm_n} \end{array} \\ \begin{bmatrix} \boldsymbol{W}^{11} & \cdots & \boldsymbol{W}^{i1} & \cdots & \boldsymbol{W}^{n1} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}^{1j} & \cdots & \boldsymbol{W}^{ij} & \cdots & \boldsymbol{W}^{nj} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}^{1n} & \cdots & \boldsymbol{W}^{in} & \cdots & \boldsymbol{W}^{nn} \end{bmatrix} \end{array} \begin{array}{l} \\ D_1 \\ \vdots \\ D_j \\ \vdots \\ D_n \end{array} \begin{array}{l} \begin{array}{l} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \end{array} \\ \begin{array}{l} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jm_j} \end{array} \\ \begin{array}{l} c_{n1} \\ c_{n2} \\ \vdots \\ c_{nm_n} \end{array} \end{array} \tag{6}$$

Step 6: Obtain the normalized $\boldsymbol{T}_D^{\alpha}$ from total degree of influence-related matrix $\boldsymbol{T}_D$ by dimension. The matrix $\boldsymbol{T}_D = \left[t_{ij}^{D}\right]_{nxn}$ by dimension can first be derived based on the dimensions (or clusters) from $\boldsymbol{T}_C$. Then, the weights versus

each criterion in dimension can be derived by the total influence-related matrix $\boldsymbol{T}_D$ for normalized the un-weighted super-matrix $\boldsymbol{W}$.

$$\boldsymbol{T}_D = \begin{bmatrix} t_{11}^{D_{11}} & \cdots & t_{1j}^{D_{1j}} & \cdots & t_{1m}^{D_{1m}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{i1}^{D_{i1}} & \cdots & t_{ij}^{D_{ij}} & \cdots & t_{im}^{D_{im}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{m1}^{D_{m1}} & \cdots & t_{mj}^{D_{mj}} & \cdots & t_{mm}^{D_{mm}} \end{bmatrix} \begin{matrix} \longrightarrow d_1 = \sum_{j=1}^{m} t_{1j}^{D_{1j}} \\ \\ \longrightarrow d_i = \sum_{j=1}^{m} t_{ij}^{D_{ij}} \\ \\ \longrightarrow d_m = \sum_{j=1}^{m} t_{mj}^{D_{mj}} \end{matrix} \quad , d_i = \sum_{j=1}^{m} t_{ij}^{D_{ij}}, i = 1,...,m \tag{7}$$

The total influence-related matrix $\boldsymbol{T}_D$ based on Eq. (7) can be normalized as follow Eq. (8).

$$\boldsymbol{T}_D^{\alpha} = \begin{bmatrix} t_{11}^{D_{11}} / d_1 & \cdots & t_{1j}^{D_{1j}} / d_1 & \cdots & t_{1m}^{D_{1m}} / d_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{i1}^{D_{i1}} / d_i & \cdots & t_{ij}^{D_{ij}} / d_i & \cdots & t_{im}^{D_{im}} / d_i \\ \vdots & \vdots & & \vdots & \vdots \\ t_{m1}^{D_{m1}} / d_m & \cdots & t_{mj}^{D_{mj}} / d_m & \cdots & t_{mm}^{D_{mm}} / d_m \end{bmatrix} = \begin{bmatrix} t_D^{\alpha 11} & \cdots & t_D^{\alpha 1j} & \cdots & t_D^{\alpha 1n} \\ \vdots & & \vdots & & \vdots \\ t_D^{\alpha i1} & \cdots & t_D^{\alpha ij} & \cdots & t_D^{\alpha in} \\ \vdots & & \vdots & & \vdots \\ t_D^{\alpha n1} & \cdots & t_D^{\alpha nj} & \cdots & t_D^{\alpha nn} \end{bmatrix} \tag{8}$$

where $\alpha_{ij}^{\alpha ij} = t_{ij}^{D_{ij}} / d_i$. The unweighted supermatrix can serve as the basis for deriving the weighted supermatrix.

Step 7: Obtain the weighted super-matrix. Then, the authors normalize $\boldsymbol{T}_D^{\alpha}$ multiply into the unweighted supermatrix to obtain the weighted supermatrix as shown in Eq. (15).

$$\boldsymbol{W}^{\alpha} = \boldsymbol{T}_D^{\alpha} \times \boldsymbol{W} = \begin{bmatrix} t_D^{\alpha 11} \times \boldsymbol{W}^{11} & \cdots & t_D^{\alpha i1} \times \boldsymbol{W}^{i1} & \cdots & t_D^{\alpha n1} \times \boldsymbol{W}^{n1} \\ \vdots & & \vdots & & \vdots \\ t_D^{\alpha 1j} \times \boldsymbol{W}^{1j} & \cdots & t_D^{\alpha ij} \times \boldsymbol{W}^{ij} & \cdots & t_D^{\alpha nj} \times \boldsymbol{W}^{nj} \\ \vdots & & \vdots & & \vdots \\ t_D^{\alpha 1n} \times \boldsymbol{W}^{1n} & \cdots & t_D^{\alpha in} \times \boldsymbol{W}^{in} & \cdots & t_D^{\alpha nn} \times \boldsymbol{W}^{nn} \end{bmatrix} \tag{9}$$

Step 8: Obtain the influential weights using DEMATEL based network process (DNP). The final step is to obtain the limit weighted supermatrix $\boldsymbol{W}^{\alpha}$. The weighted supermatrix $\boldsymbol{W}^{\alpha}$ is multiplied by itself multiple times to obtain the limit supermatrix (the concept based on Markov Chain). Then, the influential weights of each criterion can be obtained by $\lim_{g \to \infty} (\boldsymbol{W}^{\alpha})^g$; in other word, the influential weights of DNP can be obtained and denoted the limit supermatrix $\boldsymbol{W}^{\alpha}$ with power g (g representing any number for power); we call this process as DNP.

4 Empirical Study: Case of Smart TV

The Smart TV is an innovative product for consumers. To promote and design the Smart TV, designers and sales have to realize the consumers' technology acceptance of the Smart TV. This research applied the LUM and DNP methods to predict the consumers' technology acceptance. In order to verify the framework being mentioned in Section 3 and demonstrate the efficiency of the LUM in differentiating the casual relationships being derived based on the lead users, an empirical study based on the research results by surveying three experts now serving in the world's leading IT firms providing Smart TV related products.

At first, the factors influencing the acceptance of Smart TVs were collected based on literature review. These factors include (1) the perceived usefulness (PU), (2) the perceived ease of use (PEU), (3) the subjective norm (SN), (4) the perceived behavioral control (PBC), (5) the attitude (ATT), (6) the behavioral intention (BI), (7) the actual system use where (B) refers to the actual behaviors of users (Davis 1986).

In order to derive the causal structure based on the opinions of the lead users', three experts now serving in the world's leading IT firms were invited. Then, the causal structure can be derived by using the DNP. After the derivation of the total influence matrix, the casual relationship can be derived by setting 0.753 as the threshold. Based on the empirical study results, the DNP, PU, PEU, BI, ATT, PBC, SN and B can serve as the factors for predicting the acceptance of the Smart TVs. However, there is no influence between the SN and other criteria. Consequently, the actual purchase can't be predicted by the PBC. The casual structure being derived by using the DNP is demonstrated in Fig.4.

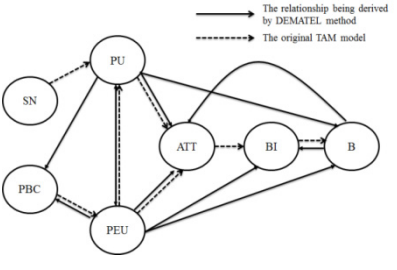


Fig. 4 The Causal Relationship of Criteria

Further, the DNP is applied to derive the weights versus each factor and the weights are shown in Table 1. According to the weights being derived, the B, ATT, PEU and PBC are important factors based on the viewpoint of the lead users while the SN is regarded as non-vital.

Table 1 The Weight of Criteria

Criteria	PU	PEU	ATT	BI	SN	PBC	B
Weight	0.134	0.146	0.151	0.139	0.127	0.144	0.159
Ranking	6	3	2	5	7	4	1

5 Discussions and Managerial Implications

Regarding to the managerial implication, both the casual relationship structure and strength of influences between the factors are different from the aspects of the lead users. From the aspect of the lead users, the B, ATT, PEU and PBC are recognized as significant factors for influencing customers' acceptance of the Smart TVs. However, there is no causal relationship between the SN and other criterion, the SN is not suitable to predict the mass customers' preference. On the other hand, the actual usage will influence and be influenced significantly. The PU and PEU will also influence the actual usage directly. Further, the ATT and the BI will be influenced by the actual usage. Regarding to the importance of predicting acceptance factors, the PEU and the PU influence other factors the most.

The DNP is applied to derive the weights versus each factor, which are 0.134. 0.146, 0.151, 0.139, 0.127, 0.144, and 0.159 for the PU, PEU, ATT, BI, SN, PBC, and B, respectively. Regarding to the managerial implication, actual usage is the most important criterion from lead users' opinions. The result can conclude that the Smart TV companies should promote the Smart TVs by actual trial. After the customers' actual trial, their attitude and behavior intention will be change. Then these changes will increase the willingness to buy the Smart TVs.

6 Conclusions

The consumer behavior prediction of Smart TVs is a difficult and indispensable task due to the fast emerging technology and severe competitions of the vendors. The DNP can be applied to solve the problem without defined casual structure. This research attempted to predict the acceptance of mass customers' by using LUM and novel DNP method being proposed by Prof. Gwo-Hshiung Tzeng. According to the analytic results, the perceived ease of use, attitude and the actual usage are important criteria for predicting the acceptance of mass customers.

References

Ajzen, I.: Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. Advances in Experimental Social Psychology 20, 1–63 (1987)

Ajzen, I.: Attitudes, personality, and behavior. Open University Press, Milton Keynes (1989)

Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)

Bagozzi, R.P.: Attitudes, intentions, and behavior: A test of some key hypotheses. Journal of Personality and Social Psychology 41(4), 607–627 (1981)

Beck, H., Mylonas, A., Rasmussen, R.: Business Communication and Technologies in a Changing World. Macmillan Education Australia (2009)

Chiu, Y.J., Chen, H.C., Shyu, J.Z., Tzeng, G.H.: Marketing strategy based on customer behavior for the LCD-TV. International Journal of Management and Decision Making 7(2/3), 143–165 (2006)

Davis, F.D.: A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral dissertation (1986)

Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. Management Science 35(8), 982–1003 (1989)

Fishbein, M., Ajzen, I.: Attitudes toward objects as predictors of single and multiple behavioral criteria. Psychological Review 81(1), 59–74 (1974)

Gabus, A., Fontela, E.: World Problems, an Invitation to Further Thought Within the Framework of DEMATEL. Batelle Geneva Research Center, Geneva (1972)

Ho, W.R., Tsai, C.L., Tzeng, G.H., Fang, S.K.: Combined DEMATEL technique with a novel MCDM model for exploring portfolio selection based on CAPM. Expert Systems with Applications 38(1), 16–25 (2011)

Hippel, E.V.: Lead Users: A Source of Novel Product Concepts. Management Science 32(7), 791–805 (1986)

Hu, L.T., Bentler, P.M.: Cut off criteria for fit indexes in covariance. Structural Equation Modeling 6(1), 1–55 (1999)

Huang, C.Y., Tzeng, G.H.: Reconfiguring the Innovation Policy Portfolios for Taiwan's SIP Mall Industry. Technovation 27(12), 744–765 (2007)

Huang, C.Y., Hong, Y.H., Tzeng, G.H.: Assessment of the Appropriate Fuel Cell Technology for the Next Generation Hybrid Power Automobiles. Journal of Advanced Computational Intelligence and Intelligent Informatics (2011) (forthcoming)

Huang, C.Y., Tzeng, G.H., Ho, W.R.: System on chip design service e-business value maximization through a novel MCDM framework. Expert Systems with Applications 38(7), 7947–7962 (2011)

Jeng, D.J.F., Tzeng, G.H.: Social influence on the use of clinical decision support systems: revisiting the unified theory of acceptance and use of technology by the fuzzy DEMATEL technique. Computers & Industrial Engineering (in press, 2012)

Kuan, M.J., Tzeng, G.H., Hsiang, C.C.: Exploring the quality assessment system for new product development process by combining DANP with MCDM model. International Journal of Innovative Computing, Information and Control (special issue) 8(8) (in press, 2012)

Liou, J.J.H., Tzeng, G.H.: A hybrid ANP model in fuzzy environments for strategic alliance partner selection in the airline industry. Applied Soft Computing 11(4), 3515–3524 (2011)

Liou, J.J.H., Tzeng, G.H., Hsu, C.C., Yeh, W.C.: Reply to comment on using a modified grey relation method for improving airline service quality. Tourism Management (in press, 2012)

Liu, C.H., Gwo-Hshiung Tzeng, G.H., Lee, M.H.: Improving tourism policy implementation - the use of hybrid MCDM models. Tourism Management 33(2), 239–488 (2012)

Liu, C.H., Tzeng, G.H., Lee, M.H.: Strategies for improving cruise product sales in the travel agency- using hybrid MCDM models. The Service Industry Journal (in press, 2012)

Liou, J.J.H., Tsai, C.Y., Lin, R.H., Tzeng, G.H.: A modified VIKOR multiple-criteria decision method for improving domestic airlines service quality. Journal of Air Transport Management 17(2), 57–61 (2011)

Olson, J.M., Zanna, M.P.: Attitudes and attitude change. Annual Review of Psychology 44, 117–154 (1993)

Ou Yang, Y.P., Shieh, H.M., Tzeng, G.H.: A Ikor Technique Based on Dematel and ANP for Information Security Risk Control Assessment. Information Sciences (in press, 2012)

Patrick, M.: Make Your TV Smarter. PC World 29(8), 58–65 (2011)

Rogers, E.M.: Diffusion of innovation (1962) (accessed)

Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

Saaty, T.L.: Decision Making With Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburgh (1996)

Schumacker, R.E., Lomax, R.G.: A Beginner's Guide to Structural Equation Modeling. Lawrence Erlbaum Associates, Publishers, Mahwah (1996)

Shen, Y.C., Lin, G.T.R., Tzeng, G.H.: Combined DEMATEL techniques with novel MCDM for the organic light emitting diode technology selection. Expert Systems with Applications 38(3), 1468–1481 (2011)

Tzeng, G.-H., Huang, C.-Y.: Combined DEMATEL technique with hybrid MCDM methods for creating the aspired intelligent global manufacturing & logistics systems. Annals of Operations Research, 1–32 (in press, 2012)

Urban, G., Von Hippel, E.: Lead user analyses for the development of new industrial Products. Management Science 34(5), 569–582 (1988)

Wang, Y.L., Tzeng, G.H.: Brand marketing for creating brand value based on a MCDM model combining DEMATEL with ANP and VIKOR methods. Expert Systems with Applications 39(5), 5600–5615 (2012)

Wei, P.L., Huang, J.H., Tzeng, G.H., Wu, S.I.: Causal modeling of Web-Advertising Effects by improving SEM based on DEMATEL technique. Information Technology & Decision Making 9(5), 799–829 (2010)

Yang, J.L., Tzeng, G.H.: An integrated MCDM technique combined with DEMATEL for a novel cluster-weighted with ANP method. Expert Systems with Applications 38(3), 1417–1424 (2011)

Yoo, C.S.: Product Life Cycle Theory and the Maturation of the Internet. Northwestern University Law Review 104(2) (2010)

Designing a Rule Based Expert System for Selecting Pavement Rehabilitation Treatments Using Rough Set Theory

Ehsan Akhlaghi and Farshid Abdi

Abstract. Rule based expert systems provide a suitable method for capturing expert knowledge and formulating the acquired knowledge as rules. In the field of pavement rehabilitation, the decision making process includes choosing the best rehabilitation or maintenance options according to the values of different pavement conditions and indexes. The decision about the same road pavement may vary significantly from one expert to another because of their different experiences and attitudes. Thus, using an expert system which is composed of rules extracted from experts consensus could be viewed as a beneficial pavement management decision support tool. Rough Set theory is proven to be appropriate for processing qualitative information that is difficult to analyze by standard statistical techniques. The current research uses rough set theory in order to derive decision rules from diverse opinions of experts in selecting pavement rehabilitation treatments.

Keywords: Rehabilitation, Expert System, Rough Set Theory, Rule Induction.

1 Introduction

Road pavement information management is a process which requires comprehensive information from the current condition of the road, and accomplishes based on engineering judgements concerning measured road indexes.

Ehsan Akhlaghi
Faculty of Industrial Engineering, Islamic Azad University, Tehran South Branch
e-mail: st_e_akhlaghi@azad.ac.ir

Farshid Abdi
Faculty of Industrial Engineering, Islamic Azad University, Tehran South Branch
e-mail: farshidabdi@azad.ac.ir

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 471–480.
springerlink.com

A pavement management system (PMS) provides effective tools and methods that help decision makers to follow efficient strategies to construct and maintain a pavement network over a planning horizon. Choosing a suitable pavement rehabilitation treatment, according to the existing information about a pavement network is a major challenge in front of decision makers. The challenge appears when experts judgements referring the same information are different as a result of their different level of knowledge and experience. Moreover, the experts are tended to select the rehabilitation options on the basis of their previous decisions in similar cases and may not have enough specialities to assess all the possible options. As the pavement rehabilitation treatments are very expensive in nature, the result of adopting improper decisions causes to huge waste of asset and time. In fact, it is vital to adopt the most cost effective method of maintenance and repair in order to achieve acceptable return in investment. Therefore, for choosing feasible rehabilitation treatments, other factors like available assets, effectiveness of the corrective strategies, and impact of the action on the system performance, should be considered too.

By emerging expert systems technology as a part of artificial intelligence which can be used to substitute for human expertise in different fields, researches could find a good potential to apply them in the field of pavement management and maintenance. Expert systems are powerful tools that able to embody the knowledge, experience, and judgement of expert pavement engineer by first identifying a main feasible rehabilitation approach and to provide the local engineering user with an interactive analysis and design tool for development of pavement rehabilitation strategies. [14]

Rose, Sceptre, Preserver, Erasme, Expear and Pavement Expert are several examples of late 1980s developed expert systems in pavement management field. They were designed especially for inferring highways rehabilitation solutions. [1, 2, 3, 4] Later in early 1990s an expert system called Paver evolved which was also applicable to airport networks. [5]

The aim of this study is to gather knowledge from experts in the field of pavement rehabilitation and to induce rules from a data set of their knowledge and discover data regularities. To achieve this purpose, we constitute a panel of field experts who use their knowledge, judgement and experience to make interferences. It is hard to reach the members of the panel in local agencies and their necessary knowledge in selecting the proper treatments is difficult to transfer to the less experienced engineers and when they get retired, the knowledge and experience may be lost. Thus, it seems to be needful to capture the experts knowledge as much as possible. We did this through conducting knowledge acquisition surveys.

2 Rough Set Theory

Rough set theory (RST), firstly introduced by a Polish computer scientist Pawlak [8], has proved to be a powerful tool for dealing with uncertainty and has been applied to data reduction, rule extraction, data mining and granularity computation [12]. The rough set concept can be of some importance, primarily in some branches of artificial intelligence, such as inductive reasoning, automatic classification, pattern recognition, learning algorithms, etc [8]. The essence of rough set theory is to reduce a given decision table to the decision table small enough so that decision rules can be directly extracted [3]. Thus, we can consider rough set theory as a data reduction technique in data mining. The three kinds of data reduction using rough set theory are row reduction, value reduction and column reduction. Row reduction is to summarize similar rows in a table, value reduction tries to simplify decision rules and column reduction is to achieve more relevant columns.

A rough set is an approximation of a crisp set in terms of a pair of sets which give the lower and the upper approximation of the original set. In rough set theory, an information system is described by $I = (U, A)$, where $U = \{x_1, x_2, ..., x_n\}$ is a non-empty finite set of objects (universe) and $A = \{a_1, a_2, ..., a_m\}$ is a non-empty finite set of attributes (features) such that $a : U \rightarrow V_a$ for every $a \in A$. The theory denotes V_a as the domain of the attribute a or the set of values that attribute a may take. A decision or data table, assigns a value from V_a to each attribute a and object x in the universe using an information function $f : U \times A \rightarrow V$ such that $f(x_i, a) \in V_a$ for each $a \in A$ and $x_i \in U$. The set A is usually divided into set C of condition attributes and a set D of decision attributes $A = C \cup D$.

In RST, the objects in universe U can be described by various attributes in A. When two different objects have the same values of conditional attributes, then these two objects are classified as indiscernible in the information system I, thus their relationship leads to indiscernibility relation. In mathematical terms, an indiscernibility relation $IND(B)$ generated by attribute subset $B \subset A$ on U, is defined as follows: [11]

$$IND(B) = \{(x_i, x_j) \in U \times U | f(x_i, a) = f(x_j, a), \forall a \in B\}$$

For example, Table 1 shows a decision table with five objects that are characterized with four condition attributes (F1, F2, F3 and F4) and one decision attribute (O).

Reduct is a fundamental construct in rough set theory which is defined as a minimal sufficient subset of attributes ($Reduct \subseteq A$) that produces the same categorization of objects as the collection of all attributes. To produce reducts, the algorithm shown in Fig. 1 can be used, where m is the number of attributes. The generated reducts based on the algorithm for the objects in Table 1 is shown in Table 2.

Table 1 Sample decision table

Object No.	F1	F2	F3	F4	O
1	1	0	2	0	0
2	1	0	0	1	1
3	0	0	3	1	0
4	1	1	2	0	2
5	0	0	1	0	0

Table 2 Reducts generated for objects in Table 1

Reduct No.	Object No.	F1	F2	F3	F4	O
1	1	x	0	x	0	0
2	1	x	0	2	x	0
3	2	x	x	0	x	1
4	3	x	x	3	x	0
5	3	0	x	x	x	0
6	4	x	1	x	x	2
7	5	x	x	1	x	0
8	5	0	x	x	x	0

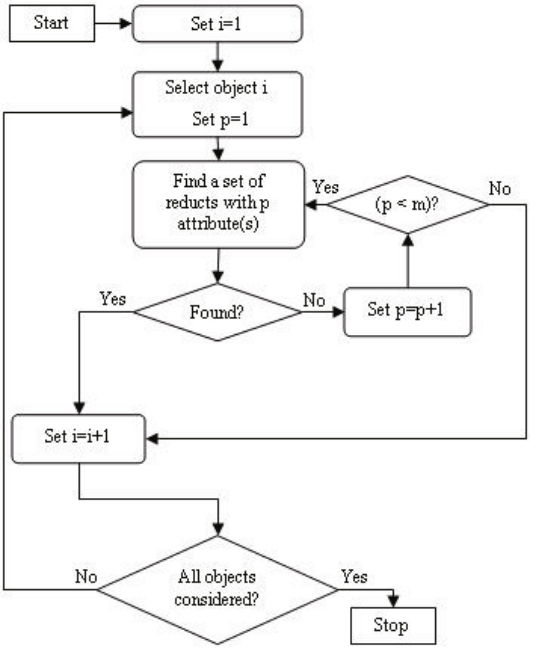


Fig. 1 Reduct generation flow chart

In Table 2, sign "x" in each row implies that the attribute is not considered to determine the output of the object in that row. In other words it is a dontcare value in making the current reduct. The values 0, 1, 2 and 3 may refer to linguistic terms like "rare", "low", "medium" and "high", respectively.

For example, if we follow the algorithm from the first step, we select the first object and try to find a set of reducts with only one attribute. It is obvious that there is no single attribute value for condition attributes (F1, F2, F3 or F4) which can solely determine the decision attribute (O) with the value of zero. Indeed, if we select the value of F1 in the first object which has O=0, there are two other objects (object 2 and 4) who have the same value for F1 and different values for O and this violates the indiscernibility relations condition. The same situation exists for single attribute values of F2, F3 and F4. Following the algorithm, we set p=2 and as the number of attributes is bigger than 2, we try to a find set of reducts using two attributes. This will result in finding F2=0 and F4=0, and also F2=0 and F3=2, which will imply the indiscernibility condition together. Because we have found reducts, we will not continue to increment p and will go to next object. The same method will be adopted to extract the whole reducts from Table 1.

3 Road Pavement Indexes

Road pavement indexes are specific factors which should be measured in a pavement sector in order to understand the current condition of the pavement. By assessing the road pavement indexes, an expert can recognize the major deficiencies of a road and prescribe proper rehabilitation or maintenance treatments to improve the quality of the pavement and prolong the lifetime of the road. Ideally, the acceptable value of indexes shows that no rehabilitation or maintenance action is necessary to be made on the road, but the out of range values of indexes exert that rehabilitation or maintenance actions should be made.

In the current study, we have selected 13 determinant road pavement indexes which most of road experts refer to for analyzing the status of a road. The list of these indexes and a brief description to them is displayed in Table 3. [13].

4 Rehabilitation and Maintenance Treatments

The road experts select appropriate rehabilitation and maintenance treatments according to situations of the road pavement. Indeed, the road pavement indexes reveal the current state of the road and the experts decide to choose expedient actions based on their experience and knowledge. The treatments list, showed in Table 4, describes 12 prevalent items which are mostly used by road experts in this research.

Table 3 Determinant road pavement indexes

No.	Index Name	Description
1	PCI	Indicates the condition of a roadway
2	IRI	Measures pavement smoothness
3	SN	Indicative of the total pavement thickness
4	Alligator Cracks	Interconnected cracks caused by fatigue failure of asphalt concrete surface under repeated traffic loading
5	Block Cracks	Interconnected cracks that divide the pavement up into rectangular pieces
6	Longitudinal Cracks	Cracks parallel to the pavement's centerline or lay-down direction
7	Transverse Cracks	Cracks extend across the pavement
8	Cracks Depth	The depth of the cracks of asphalt pavement
9	Rutting	The permanent deformation of pavement layers which can accumulate over time
10	Ravelling	Disintegration of the pavement from the surface downward
11	Bleeding	Occurs in bituminous pavement when a film of asphalt binder appears on road surface
12	Weathering	The hardening and aging of the asphalt binder
13	AADT	Average annual daily traffic

5 Knowledge Acquisition

In this research, knowledge acquisition phase has performed by conducting a survey. The survey has simply composed by a table which its columns represent pavement indexes as condition attributes and the treatment as decision attribute, in rough set terms. Complete explanation about each pavement index and rehabilitation treatment is offered to make the same sense for survey participants.

Each pavement index (condition attribute) has been classified into five levels: "very low", "low", "middle", "high", and "very high", and represented by "1", "2", "3", "4" and "5". Value "x" means that the index value is not important for choosing the treatment. For instance, if IRI index (F2) = 3, means the IRI value is in the middle of its range; then F5 = 1, means the number of block cracks in the road section is high. For the rehabilitation treatment (O, which is the decision attribute), we also classified it into the number of corresponding item in treatment alternatives list (the first column of Table 4). Table 5 is a typical survey form.

If the expert selects more than one option for the treatment (O), we translate this choice to multiple rows with the same index values and each of the selected treatments as decision attribute for them. For example, the three last rows of Table 5 have exactly same values for indexes and different treatment

Table 4 List of alternative treatments

No.	Treatment	Description
1	Full Pavement Replacement	The removal and replacement of the old pavement
2	Thin Overlay	Overlaying the pavement surface with a thin layer of asphalt
3	Thick Overlay	Overlaying the pavement surface with a thick layer of asphalt
4	Mill and Overlay	Grinding to smooth a pavement and then using overlays
5	Cold Recycling	Using cold-mix asphalt to recycle the pavement
6	Hot Recycling	Using hot-mix asphalt to recycle the pavement
7	Micro-surfacing	A polymer-modified cold-mix paving system
8	White-topping	Covering of an existing asphalt pavement with a layer of Portland cement concrete
9	Crack Sealing	Filling or sealing pavement cracks to prevent water from entering the base and sub-base
10	Chip Seal	Extends pavement life and provides a good driving surface
11	Fog Seal	Sealing the voids and prevent further deterioration
12	No Action	When the pavement indexes are in their appropriate range

Table 5 A typical answered survey form

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	O
1	x	x	2	x	x	x	x	1	5	x	x	x	3
1	x	4	2	1	1	1	x	x	x	x	x	x	4
x	x	3	3	x	x	2	x	5	5	5	x	x	11
x	x	1	1	1	x	5	x	x	x	2	x	x	8
1	x	5	1	1	1	x	x	x	x	x	x	x	3
1	x	5	1	1	1	x	x	x	x	x	x	x	9
1	x	5	1	1	1	x	x	x	x	x	x	x	8

values. This means that if F1 (PCI), F4 (number of Alligator Cracks), F5 (number of Block Cracks), F6 (number of Longitudinal Cracks) are very low and F3 (SN) is very high, the suitable treatment option by the expert's point of view, is thick overlay, white-topping and crack sealing. Experts can also fill multiple rows with different index values and the same treatment option.

We can define a conflict between experts opinions when at least two different experts suggest same values for conditional attributes (F1, ..., F13) and different values for decision attribute (O). In the case of arising conflicts, according to [9] and [10], we may define a basic relation R^- over U^2 called conflict relation and defined as follows:

$$R^-(x,y) \quad iff \quad a_i(x) = a_i(y) \quad and \quad a_m(x) \neq a_m(y), \qquad i = \{1, ..., m-1\}$$

In the conflict relation, m represents the number of attributes (including a single decision attribute as the m'th attribute) and $a_i(x)$ and $a_i(y)$ are the values assigned to i'th attribute by experts x and y, respectively. Using conflict relation we can shape conflict graphs and in order to change the conflict situation we need negotiations between experts.

6 Rule Validation

Validation process is critical to the design and implementation of decision-making expert systems and ascertains what the system knows, what it incorrectly knows, and what it does not know; that is, validation ascertains accuracy and completeness. [7]

In our decision table, we randomly divide data into the training and testing data sets. According to [6], we can split data via the bootstrapping method with the following ratios: 0.6232 for training set and 0.368 for testing set. The training data set is used to build the model and extract rules and the testing data is applied to detect over fitting of the modelling tools.

Table 6 Determinant road pavement indexes

No.	Rule	Supporting Objects	Accuracy
1	If (PCI is middle) and (number of Alligator Cracks is Low) and (number of Block Cracks is Low) and (number of Longitudinal Cracks is Middle) and (number of Transverse Cracks is Middle) and (SN is Low) and (AADT is Low) THEN Treatment is Thick Overlay and Crack Sealing.	29	91.3%
2	If (PCI is Very High) and (number of Block Cracks is Low) and (Weathering is High) THEN Treatment is Fog Seal.	22	86.8%
3	If (PCI is Low) and (number of Block Cracks is Middle) and (number of Alligator Cracks is Very Low) and (IRI is High) and (SN is Very Low) THEN Treatment is Crack Sealing and Thin Overlay and Whitetopping.	19	85.18%

In order to estimate the accuracy of the rules, we should compare each derived rule with each object from the testing data set and calculate the number of objects that match with the rule. The accuracy of each rule could be calculated using the total matched objects divided by the summation of the total correctly matched objects and the total incorrectly matched objects.

By comparing the resulted accuracy of a rule with a predefined threshold, we may approve or reject the rule. Note that in our decision table, an incorrectly matched object means that the object the identical value of pavement indexes with the rule, yet the treatment option is different from the rule.

7 Results

Twenty five experts from credible universities and companies participated in the survey. Aggregating all survey forms, we constructed the information system and using fuzzy indiscernibility relation, the corresponding reduct set was found. For producing reduct set we used Roughiran software which is developed by Pishtazane Asre Kavosh company in Iran. The software generates reduct set in the form of a SQL Server database table. The software also computes the number of supporting objects for each reduct. The supporting objects for a reduct are the rows in information system which the reduct has derived from them. So, the number of supporting objects could be considered as an index of consensus among experts opinions.

Thirty eight rules were induced from the reduct set. According to the ratio 0.632 for training data set and 0.368 for testing data set, 24 reducts were randomly selected for training and 14 reducts for testing. Table 6 displays the three most accurate rules. The rules are ordered by the number of supporting objects.

8 Conclusion

Obtaining decision rules from various opinions of experts in the field of pavement rehabilitation treatment has been presented in the paper. The rules inducted from calculated reduct set, could be considered as the basis of a rule based expert system. The similar method of rule generation can be carried out including a bigger size panel of experts to obtain more accurate results. Also, as the level of knowledge and experience of experts may vary, the researches can be conducted using different weightings for experts decisions.

Acknowledgements. The authors would like to thank the staff of "Tadbir Foroudrah" and "Pishtazane Asre Kavosh" company for their assistance. This work is supported by "National Elites Foundation of Iran" and "Islamic Azad University".

References

1. Allez, F., Dauzats, M., Joubert, P., Labat, G.P., Pusselli, M.: ERASME: An Expert System for Pavement Maintenance. Transportation Research Record 1205, 1–5 (1988)

2. Al-Shawi, M.A., Cabrera, J.G., Watson, A.S.: Pavement Expert: An Expert to Assist in the Evaluation of Concrete Pavements. In: Proceeding of Transportation and Planning Meeting, Leeds, England, Leeds, England, p. 293 (1989)
3. Haas, R., Shen, H.: PRESERVER: A Knowledge-Based Pavement Maintenance Consulting Program, Advanced Development Department Computing Devices Company (1989)
4. Hajek, J.J., Haas, R., Chong, G.J., Phang, W.A.: ROSE: A knowledge-based expert system for routing and sealing. In: Proceeding of the 2nd North American Pavement Management Conference, Toronto, Canada, pp. 2.301–2.341 (1987)
5. Ismail, N., Ismail, A., Rahmat, R.A.: Development of Expert System for Airport Pavement Maintenance and Rehabilitation. European Journal of Scientific Research 35(1), 121–129 (2009) ISSN 1450-216X
6. Kusiak, A.: Rough set theory: A data mining tool for semiconductor manufacturing. IEEE Transactions on Electronics Packaging Manufacturing 24, 44–50 (2001)
7. Oleary, D.E.: Methods of Validating Expert Systems. Interfaces 18(6), 72–79 (1988)
8. Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11, 341–356 (1982)
9. Pawlak, Z.: On Conflicts. International Journal of Man-Machine Studies 21, 127–134 (1984)
10. Pawlak, Z.: Conflicts and Negotations. In: Wang, G.-Y., Peters, J.F., Skowron, A., Yao, Y. (eds.) RSKT 2006. LNCS (LNAI), vol. 4062, pp. 12–27. Springer, Heidelberg (2006)
11. Pawlak, Z., Skowron, A.: Rudiments of rough sets. Information Sciences 177(1), 3–27 (2007)
12. Pawlak, Z., Skowron, A.: Rough sets: Some extensions. Information Sciences 177(1), 28–40 (2007)
13. Skok, G., Westover, T., et al.: Pavement Rehabilitation Selection, Minnesota Department of Transportation, USA (2008)
14. Sundin, S., Corinne, B.L.: Artificial Intelligence-Based Decision Support Technologies in Pavement Management. Computer Aided Civil and Infrastructure Engineering 16, 143–157 (2001)

Detection of Comoving Groups in a Financial Market

Takeo Yoshikawa, Takashi Iino, and Hiroshi Iyetomi

Abstract. We reported a correlation structure hidden in the Tokyo Stock Exchange (TSE) market at KES-IDT2011. By regarding the TSE market as a network (stocks and correlation coefficients correspond to nodes and weights of links between nodes, respectively) and minimizing the Frustration among nodes, the stocks were decomposed into four comoving groups forming communities. Three of them are strongly anticorrelated to each other, and the remainder is comparatively neutral to the rest of the communities. In this paper we further extend the previous work to detect tightly-coupled groups within the communities; "Hamiltonian" is used instead of the Frustration. The Hamiltonian has two parameters which control degree of strength for correlations to be extracted. It is found that six sectors (Electric Appliance, Banks, Electric Power & Supply, Information & Communication, Securities & Commodity Futures, and Insurance) form strong cores in the communities.

1 Introduction

Minimizing the Hamiltonian [1] is used to detect communities of networks with positive links as well as negative links. The Hamiltonian has two parameters γ^+ and γ^-. Changing the values of these, it becomes possible to take arbitrary correlation strength for detecting groups. In the previous work [2], we used the Frustration, which is the Hamiltonian with $\gamma^{\pm} = 0$, to detect community structure. In this paper, we set the values of $\gamma^+ = 2.0$, and $\gamma^- = 0.0$ to extract more strongly correlated groups.

Takeo Yoshikawa
Graduate School of Science and Technology, Niigata University, Ikarashi, Niigata 950-2181, Japan
e-mail: tko@phys.sc.niigata-u.ac.jp

Takashi Iino · Hiroshi Iyetomi
Faculty of Science, Niigata University, Ikarashi, Niigata 950-2181, Japan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 481–486.
springerlink.com

We first filter off the noise from correlation matirx by using RMT, thus extract genuine correlation. The correlation matrix thus obtained is then regarded as an adjacency matrix to construct a stock correlation matrix. We detect groups in which nodes are strongly correlated as communities in the network.

2 Stock Correlation Network

We analyzed the daily prices of $N = 557$ stocks belonging to the Tokyo Stock Exchange (TSE) for the 10-year period 1996–2006 ($T = 2706$ daily returns). Now we consider a $N \times T$ matrix G, whose elements $\{G_{i,t}\}_{t=1,\ldots,T}$ are standardized time series of the log-returned prices changes of the stock i at time t. Then, the correlation matrix C are calculated as

$$\mathsf{C} = \frac{1}{T}\mathsf{G}\mathsf{G}^{\mathrm{T}}, \tag{1}$$

and its elements are

$$C_{ij} = \frac{1}{T}\sum_{t=1}^{T} G_{i,t}G_{i,t} \quad (i,j = 1,\ldots,N). \tag{2}$$

The correlation matrix may be decomposed [3, 4, 5] into

$$\begin{aligned}\mathsf{C} &= \mathsf{C}_{\mathrm{market}} + \mathsf{C}_{\mathrm{group}} + \mathsf{C}_{\mathrm{random}} \\ &= \lambda_1 \mathsf{u}_1 \mathsf{u}_1^{\mathrm{T}} + \sum_{i=2}^{13} \lambda_i \mathsf{u}_i \mathsf{u}_i^{\mathrm{T}} + \sum_{i=14}^{557} \lambda_i \mathsf{u}_i \mathsf{u}_i^{\mathrm{T}},\end{aligned} \tag{3}$$

where λ_i's ($\lambda_1 \geq \cdots \geq \lambda_N$) are the eigenvalues of C and u_i are the corresponding eigenvectors. If there were no correlation between stock prices, the eigenvalues shall distribute only in the range $[\lambda_-, \lambda_+]$ [6]. Hence, one can not distinguish fluctuations ascribed to u_i with $\lambda_i \leq \lambda_+$ from noise. In the TSE data, $Q \simeq 4.86$, and we obtained $\lambda_+ \simeq 2.11$. Because the 13th and 14th eigenvalues are $\lambda_{13} > \lambda_+ > \lambda_{14}$, the interaction $\sum_{i=14}^{557} \lambda_i \mathsf{u}_i \mathsf{u}_i^{\mathrm{T}}$ is regarded as being random. Furthermore, λ_1 is eight times or more as large as λ_2, and the components of u_1 are all positive. So $\lambda_1 \mathsf{u}_1 \mathsf{u}_1^{\mathrm{T}}$ indicates the market mode.

Now, we adopt $\mathsf{C}_{\mathrm{group}}$ for the adjacency matrix to construct a stock correlation network. We exclude $\mathsf{C}_{\mathrm{market}}$ because it just describes a collective motion of the whole market.

3 Community Detection

Let us consider a undirected signed graph G composed of n nodes and m links; the links can take both $+1$ and -1. The total number of positive and negative links in G are denoted as m^+ and m^-, respectively; hence $m = m^+ + m^-$. Similarly, in a

weighted graph, m^+ and m^- denote the sum of positive and negative weights of the graph, respectively. An adjacency matrix A of the network is defined as follows: if there exists a connection from node i to node j, then $A_{ij} = w_{ij}$, and $A_{ij} = 0$ otherwise, where w_{ij} is weight for the link from node i to node j. By setting $A^+_{ij} = A_{ij}$ if $A_{ij} > 0$ and $A^+_{ij} = 0$ otherwise, and $A^-_{ij} = -A_{ij}$ if $A_{ij} < 0$ and $A^-_{ij} = 0$ otherwise, we separate the positive and negative links so that $\mathsf{A} = \mathsf{A}^+ - \mathsf{A}^-$. The positive and negative weight of node i are expressed as

$$k^\pm_i = \sum_{j=1}^{n} A^\pm_{ij}. \tag{4}$$

Let us suppose that the network is partitioned into communities; the community to which node i belong is designated by σ_i. A configuration of the community assignment is thus represented by a set $\{\sigma\}$ of σ_i's.

At a given assignment $\{\sigma\}$, there exist positive links between communities and negative links within communities in networks with positive and negative links. Such a situation is measured [1] by *Frustration* expressed as

$$F(\{\sigma\}) = -\sum_{ij} A_{ij}\delta(\sigma_i, \sigma_j), \tag{5}$$

where $\delta(i, j)$ is the Kronecker delta. Minimizing the frustration $F(\{\sigma\})$ thus leads to maximizing the sum of positive weights within communities. Hence the frustration groups together all nodes into one community in networks with only positive links. For this reason, we cannot find appropriate community assignment in such networks.

Traag and Bruggeman [1] introduced an extension of modularity to accommodate more general networks. According to them, it is called Hamiltonian and defined as

$$H(\{\sigma\}) = -\sum_{ij}[A_{ij} - (\gamma^+ p^+_{ij} - \gamma^- p^-_{ij})]\delta(\sigma_i, \sigma_j), \tag{6}$$

where $\gamma^\pm$ are thresholds for clustering nodes together versus keeping them apart (Fig. 1). The Hamiltonian reduces to the frustration (5) in the limit of $\gamma^\pm = 0$ and also recovers the modularity [7] with $\gamma^\pm$ for networks with only positive links.

4 Correlation in the Stock Network

We carried out community detection in the stock correlation network within the framework of Hamiltonian with $\gamma^+ = 2.0$ and $\gamma^- = 0.0$, to detect more stronger correlated goups. We used a simulated annealing method [8] to minimize the Hamiltonian $H(\{\sigma\})$ for the network determined by $\mathsf{C}_{\mathrm{group}}$.

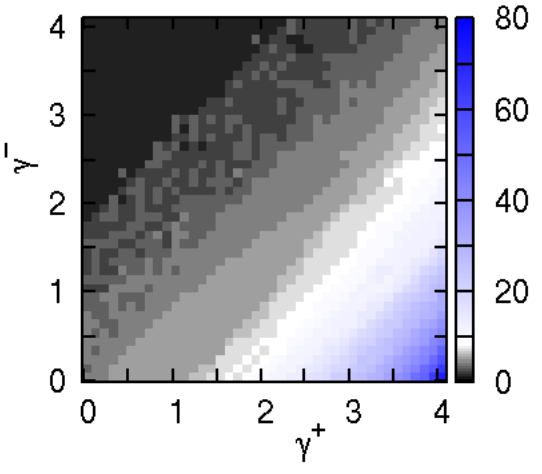


Fig. 1 $\gamma^{\pm}$dependence of the number of communities.

4.1 Correspondence for Communities

Table 1 shows the number of communities and the number of stocks of each community, obtained by the Hamiltonian minimization with two different sets of $\gamma^{\pm}$. (Call the configuration of the Hamiltonian with $\gamma^{\pm} = 0.0$ Configuration α, and that with $\gamma^{+} = 2.0$ and $\gamma^{-} = 0.0$ Configuration β.) These ten communities of configuration β are consisted of more strongly correlated stocks, because a pair of stocks is needed stronger correlation to be clustered in the Hamiltonian minimization, compared with $\gamma^{\pm} = 0.0$.

Communities of α and communities of β are mapped by containment relation of stocks as shown in Table 2.

4.2 Comparison of Correlation Structure

Minimization of Hamiltonian at $\gamma^{\pm} = 0$, or Frustration $F(\{\sigma\})$, means simply extracting groups within which nodes are connected positive links, and it is natural way against a network with positive and negative links. On the other hand, mini-

Table 1 Communitites of each configuration α and β. Configuration α is the result of Ref. [2].

Comm. No.	1	2	3	4	5	6	7	8	9	10
α	175	148	118	116	—	—	—	—	—	—
β	131	94	91	75	59	57	37	9	3	1

Table 2 Correspondence communities of α and β.
For example, 95% of stocks of comm. 3 of β are assigned comm. 4 of α.

α	Comm. 1	Comm. 2	Comm. 3	Comm. 4
	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
β	Comm. 1 (1.00)	Comm. 2 (0.99)	Comm. 4 (1.00)	Comm. 3 (0.95)

mization Hamiltonian at $\gamma^+ = 2.0$ and $\gamma^- = 0.0$ indicates more strongly correlated groups compared to Frustration minimization. Decomposition of sectors into communities is shown in Table 3. As for Electric Appliances (EA), Banks (B), Electric Power & Gas (EPG), Information & Community (IC), Securities & Comodity Futures (SCF) and Insurance (I), the groups of configuration α still remains at $\gamma^+ = 2.0$ and $\gamma^- = 0.0$, while Glass & Ceramics Products (GCP) and Real Estate (RE) become reduced in size.

Table 3 Decomposition of sectors through formation of communities. Sectors which show strong comoving behavior are listed. Fractions of stocks of individual sectors which exceed $2/3$ are highlighted in **boldface**.

	α	β	α	β	α	β	α	β
Sector \ Comm.	1	1	2	2	3	4	4	3
Electric Appliances	0.18	0.18	**0.79**	**0.74**	0.00	0.00	0.03	0.01
Banks	0.00	0.00	0.00	0.00	**1.00**	**1.00**	0.00	0.00
Glass & Ceramics Products	0.13	0.13	0.13	0.20	0.00	0.00	**0.73**	0.53
Electric Power & Gas	0.00	0.00	0.00	0.00	**1.00**	**100**	0.00	0.00
Information & Communication	0.00	0.00	**0.90**	**0.80**	0.00	0.10	0.10	0.10
Securities & Commodity Futures	0.00	0.00	**1.00**	**1.00**	0.00	0.00	0.00	0.00
Real Estate	**0.67**	0.33	0.00	0.00	0.00	0.00	0.33	0.17
Insurance	0.00	0.00	0.00	0.00	**1.00**	**1.00**	0.00	0.00

5 Summary

We revisited correlation structure in the Tokyo Stock Exchange market from a network viewpoint. The correlation matrix of stock price changes, purified by the RMT, was used to construct a network. The stock network thus constructed has weighted links and furthermore those weights can be even negative. For this reason, we employ Hamiltonian minimization to detect community of the stock correlation network, at two pairs of parameters of Hamiltonian, $\gamma^{\pm} = 0$ and, $\gamma^+ = 2.0$ and $\gamma^- = 0.0$. We found that, in TSE market, (EA), (B), (EPG), (IC), (SCF) and (I) are very strongly correlated among each sector. In addition, (EA), (IC) and (SCF) are strongly correlated and consisting the same community, similarly, (B), (EPG) and (I) are, as well. Furthermore, this procedure of community detection is applicable to

other correlation networks, analysis of time variation of correlation structures of a stock market and comparison of correlation structures between stock markets is in progress.

Acknowledgements. This work was partially supported by the Program for Promoting Methodological Innovation in Humanities and Social Sciences by Cross-Disciplinary Fusing of the Japan Society for the Promotion of Science and by the Ministry of Education, Science, Sports, and Culture, Grants-in-Aid for Scientific Research (B), Grant No. 22300080 (2010-12).

References

1. Traag, V.A., Bruggeman, J.: Phys. Rev. E 80, 036115 (2009)
2. Yoshikawa, T., Iino, T., Iyetomi, H.: Market Structure as a Network with Positively and Negatively Weighted Links. In: Watada, J., Phillips-Wren, G., Jain, L.C., Howlett, R.J. (eds.) Intelligent Decision Technologies. SIST, vol. 10, pp. 511–518. Springer, Heidelberg (2011)
3. Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L.A., Stanley, H.E.: Phys. Rev. Lett. 83, 1471 (1999)
4. Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L.A., Stanley, H.E.: Phys. Rev. E 65, 066126 (2002)
5. Kim, D.H., Jeong, H.: Phys. Rev. E 72, 046133 (2005)
6. Sengupta, A.M., Mitra, P.P.: Phys. Rev. E 60, 3389 (1999)
7. Leicht, E.A., Newman, M.E.J.: Phys. Rev. Lett. 100, 118703 (2008)
8. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Science New Series 220, 671 (1987)

Develop an Intelligence Analysis Tool for Abdominal Aortic Aneurysm

Nan-Chen Hsieh, Jui-Fa Chen, and Hsin-Che Tsai

Abstract. An Abdominal Aortic Aneurysm (AAA) is a focal dilatation at some point of the abdominal section of the aorta. In the absence of any treatment, AAA tends to grow until rupture. In this paper, we develop an Intelligence Analysis Tool to help researchers predict postoperative morbidity after AAA. The Tool includes an ensemble model, classification modules, model evaluation, and data visualization. The probabilities of complication calculated by the model of complications and a receiver operating characteristic (ROC) curve were used to evaluate the accuracy of postoperative morbidity prediction. The results show that the system proposed by this approach yields valuable qualitative and quantitative information for postoperative morbidity of Abdominal Aortic Aneurysm patients. Our System facilitates different types of users to access without to learn new data mining software.

1 Introduction

Cardiac surgery is a complex surgical operation that is indicated for patients with severe insufficiency in cardiac function. Major cardiac surgical interventions include coronary artery bypass grafting (CABG), repair of congenital heart defects, surgical treatment of atrial fibrillation, heart transplantation, repair or replacement of heart valves, aortic surgery, aneurysm repair or a combination of

Nan-Chen Hsieh
Department of Information Management, National Taipei University of Nursing and Health Sciences, Taiwan, ROC
e-mail: nchsieh@gmail.com

Jui-Fa Chen · Hsin-Che Tsai
Department of Computer Science and Information Engineering, Tamkang University, Taiwan, ROC
e-mail: alpha@mail.tku.edu.tw, shin@live.com

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 487–495.
springerlink.com

these surgical procedures. During the operation and the postoperative stay at the Intensive Care Unit (ICU) and nursing ward, there is considerable morbidity for cardiac surgery patients with postoperative complications, which results in increased hospital mortality and postoperative morbidity. Many prediction models for cardiac surgical outcome apply logistic or multivariable regression to assess preoperative risk [1-3]. An Abdominal Aortic Aneurysm in the absence of any treatment, tend to grow until rupture. There are two main techniques for Abdominal Aortic Aneurysm reparation: Open Repair (OR), which is an invasive surgical procedure, and the Endovascular Aneurysm Repair (EVAR), which is a minimally invasive procedure where a stent graft is inserted using a catheter in order to exclude the bulge from the blood circulation. Most risk assessment tools used traditional statistical method to derive prediction models for patients who were undergoing open abdominal aortic aneurysm repair, and to evaluate whether or not the surgical procedures would be successful. Similarly, postoperative morbidity is a key factor in recovery and through-put of cardiac hospital patients. Prediction of surgical mortality and postoperative morbidity is important in selecting low-risk patients for operation, and in counseling patients about the risks of undergoing surgical operation. The development of a robust prediction model can therefore both assist vascular surgeons in evaluating the expected outcome for a given patient and facilitate counseling and preoperative decision-making. Reliable and accurate prediction of operative mortality and morbidity is an essential criterion for any such risk evaluation models.

Data mining techniques are currently used in medical decision support to increase diagnostic accuracy and to provide additional knowledge to medical stuff. Their increased use provides expanded opportunities for determining the utility of medical decision making models from retrospective data. Prediction and description are two distinct concepts in data mining. For medical application, data mining not only support prediction models with satisfactory accuracy, but also needs a safety context in which decision-making activities require more explanatory support. The main distinction is that predictive data mining requires that the training dataset include an outcome variable, while descriptive data mining uses a global strategy to find the characteristics of each affinity granulation of the data. Both these data mining techniques can produce accurate, predictive and interpretable descriptive models that contribute greatly to handle medical data gathered through systematic use of clinical, laboratory, and hospital information systems. The goal of predictive data mining in clinical surgery is to derive models that can use medical data to predict patient's mortality and morbidity and thereby support clinical surgical decision-making. Predictive data mining can also aid in prognosis, diagnosis and treatment planning for surgical procedures. In contrast, descriptive data mining considers the data as affinity granulations, and aims at finding interpretable patterns and associations among data.

The use of machine learning models has become widely accepted in medical applications. Delen et al [4] developed a Web-based Decision Support Systems, which are used to build different types of prediction models (including neural networks, decision trees, ordinal logistic regression and discriminant analysis) to classify a movie. Murphy [5] presented a Clinical Decision Support System use of

Decision Tree is one of the most popularly applied methods. Artificial Neural Network (ANN) has featured in a wide range of medical applications, often with promising results. Eom et al. [6] developed a classifier ensemble-based, including Artificial Neural Network (ANN), Decision Tree (DT), and Support Vector Machine (SVM), clinical decision support system for predicting cardiovascular disease level. SVM have been successfully used in a wide variety of medical applications. Polat and Güne [7] used a least square support vector machine to assist breast cancer diagnosis. Babaoĝlu et al. [8] first used principle component analysis method to reduce data features, and acquired an optimum support vector machine model for the diagnosis of coronary artery disease. Choi [9] proposed the detection of valvular heart disorder (VHD) by wavelet packet decomposition and SVM techniques.

This study describes the development of an informative ensemble prediction model in web-based system, consisting of DT, ANN and SVM for the prediction of postoperative morbidity between preoperative variables and complication outcomes in Abdominal Aortic Aneurysm patients. For a better understanding of our study, Section 2 of this paper begins with an overview of study background and system overview in general. Section 3 describes the procedures used in this study, ensemble model for the prediction of postoperative morbidity. Section 4 discusses the experimental findings and offers observations about practical applications and directions for future research.

2 System Overview

Hospital mortality is an important clinical endpoint in major Abdominal Aortic Aneurysm. A prognostic model is often used for clinical decision-making prior to and during such surgery. The predicted outcome can be used by surgeons and patients to evaluate whether or not the surgical procedures is likely to be successful. Similarly, postoperative morbidity is a key factor in recovery and throughput of cardiac hospital patients. Prediction of surgical mortality and postoperative morbidity is important in selecting low-risk patients for operation, and in counseling patients about the risks of undergoing surgical operation. The development of a robust prediction model can therefore both assist vascular surgeons in evaluating the expected outcome for a given patient and facilitate counseling and preoperative decision-making. Reliable and accurate prediction of operative mortality and morbidity is an essential criterion for any such risk evaluation models.

The Weka workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality. It is freely available software. It is portable and platform independent because it is fully implemented in the Java programming language and thus runs on almost any modern computing platform. Weka has several standard data mining tasks, data preprocessing, clustering, classification, association, visualization, and feature selection[10]. But it is not easily accessible and inconvenient to use data analysis functions. We developed Intelligence Analysis Tool as a web application, the tool can be installed and used by

any client in the network. The corresponding algorithms in Weka library enabled data analysis tasks by interaction with the web service, and returned the analyzed results. Figure 1 shows the tool's architecture.

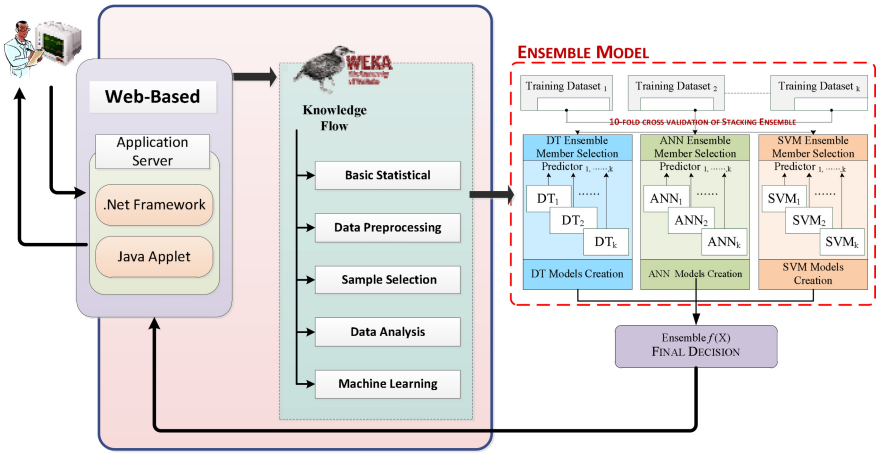


Fig. 1 The proposed architecture

The Intelligent Analysis Tool descriptive statistics functions and Data Mining functions were supported by Weka library. The Intelligent Analysis Tool descriptive statistics functions and Data Mining functions were supported by Weka library. User can choose ensemble model including DT, ANN and SVM to construct a set of hypotheses to predict risk. This study designed intelligent Analysis Tool facilitates different types of users to access functionalities in Weka without the need to learn new data mining software.

3 Method and Procedures

Abdominal Aortic Aneurysm is an enlargement that occurs in a weakened area within the largest artery in the abdomen (http://www.vascularweb.org/patients/NorthPoint/Abdominal_Aortic_Aneurysm.html). If an AAA is not treated in due time, the pressure generated by heartbeats causes the aneurysm to continuously grow larger, and the aortic wall continues to weaken. Finally, a rupture occurs that causes massive internal bleeding. The best way to prevent the high mortality associated with AAA is to find the lesion before rupture occurs. However, patients with aortic diseases are often elderly with severe co-morbidities and sometimes devastating morbidity, making them extremely challenging candidates for surgery.

3.1 Input Data

We retrospectively examined 122 consecutive patients who underwent Open Repair surgery at Taipei Veteran General Hospital, a teaching hospital in Taiwan, between 2007 and 2009. The dataset contains preoperative patient characteristics, details of the operative information, and pathological and laboratory findings from the emergency room, operating room and intensive care unit. The dataset also included length of ICU stay, variables that describe postoperative complications that frequently occur in Open Repair surgery, death during hospitalization, and time of death for patients who expired. Postoperative complication was used as the binary outcome variable of the ensemble model, and types of complications were used as subsidiary outcome variables. The original dataset contained 154 variables, but included many missing values. Preliminary inspection of the dataset showed that many variables contained missing values for at least 50% of the patients; these variables were not included in further analysis. In order to identify significant variables for use in the ensemble model, a number of criteria were employed. Variables that were subjective, ambiguous or inadequately defined were excluded; variables that were frequently incomplete were also excluded from subsequent analysis. Data collected included preoperative patient characteristics, risk factors, details of the operative information, physical characteristics of the aneurysm, postoperative physiological and laboratory findings, and postoperative complications as the outcome variable.

3.2 Ensemble Model for Prediction

DTs, ANNs, and SVMs were chosen as basis models for the ensemble model, because they represented different approaches, each of which is simple to implement and has been shown to perform well in medical applications. The rationale of employing these models is that Decision Tree have simplicity and capacity for humanly understandable inductive rules so many researchers use Decision Tree to resolve problems and error analysis. Artificial Neural Network is generally superior to conventional statistical models, and Support Vector Machine performs reasonably well in most prediction problems and can be used as a benchmark technique. Then, each individual model makes its own prediction estimating probabilities for each class. The final prediction of stacking is computed using multiple-linear regression as a Meta classifier.

The detailed configurations of each individual model are as follows: DTs with C4.5 search algorithm, ANNs are RBF neural network with radial basis functions as activation functions, SVMs use John Platt's sequential minimal optimization algorithm with logistic to the outputs for proper probability estimates.

3.3 Development the Tool as a Web-Based Application

Web-based are a common method used for achieving interoperability. The main attractions are that web-based systems are both platform and language

independent, and there exist standard protocols to facilitate the communication. A Web Services based approach generally involves setting up a web server and some proxy classes in each framework to communicate with this server. Communication is generally achieved through XML-based. .NET Framework and makes it possible both to develop .NET applications in Java, and to use existing Java API's and libraries in applications written in any .NET language. Fig. 2 shows the web-based Intelligent Data Analysis tool that centralizes to support prognosis task. Users can select the desired fields for building classifiers. The designed tool facilitates different types of users to access functionalities in Weka without the need to learn new data mining software.

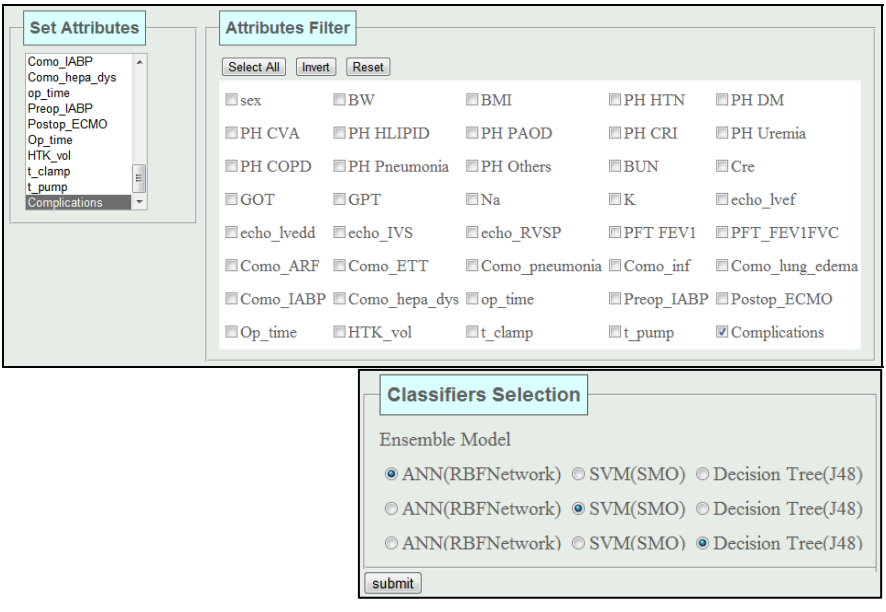
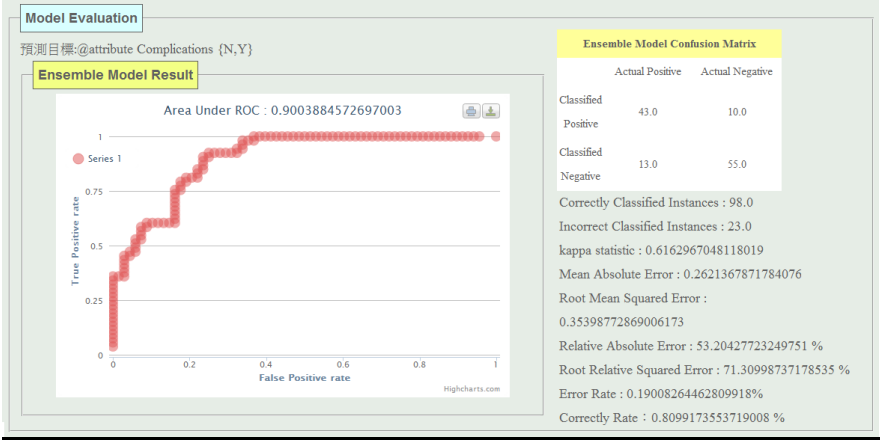


Fig. 2 Attributes filter and model selection

In this study, we intended to develop custom-designed medical data analysis components and schemata, which access the data mining suite's components through Java programming. The following functions were embedded : (1) Basic statistical functions for primary inspection of the data; (2) Visualize data exploratory functions for the interactive of interesting patterns in datasets; (3) Data sources connective and data preprocessing functions for retrieving medical data from databases, sample selection; (4) Unsupervised and supervised data analysis functions for the intelligent data analysis, includes statistical and machine learning algorithms, ex: regression analysis, clustering techniques, association rules, decision trees and support vector machines; (5) Model evaluation functions for the evaluation of built models' reliability and validity, ex: accuracy, sensitivity, specificity, ROC curves and lift chart.

The model selection scheme is a mixture of bagging and cross-validation (CV-bagging) that aims to improve the classification by combining models trained on randomly generated subsets of the entire training set. We first applied a cross validation scheme for model selection on each subset; subsequently, for the sake of simplicity and so as not to run into over-fitting problems, we combined the selected models in a uniform weights approach to reach a final decision. The concept of uniform weights voting is both appealing and simple to implement; it can generally be applied to any type of classifier without relying on specific interpretations of the output. In order to validate models, we used random subsets as cross-validation folds for reasons of simplicity. In k-fold cross validation, the dataset is partitioned into k subsets. Of these k subsets, a single subset is retained as the validation dataset, and the remaining k-1 subsets are used for training datasets. The cross validation process is then repeated k times with each of the k subsets used only once as the validation data. The k results from the folds can then be averaged to produce a single estimation. In cases where specific inside medical knowledge is not available, such a cross validation method can be used to select a classification method empirically, because it seems to be obvious that no classification method is uniformly superior. We trained several times for individual classifiers and select the best performance one as the final model. The rationale is if we trained each model with different initial conditions, we can find leverage performance for the final model. The result, a heterogeneous ensemble, allows classification methods to be used more effectively.



Models	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
10 folds CV- bagging weighted average						
DT	0.81	0.169	0.829	0.81	0.81	0.872
ANN	0.736	0.264	0.739	0.736	0.736	0.809
SVM	0.744	0.258	0.746	0.744	0.744	0.743
Ensemble	**0.81**	**0.19**	**0.812**	**0.81**	**0.81**	**0.900**

Fig. 3 The results of the prediction

To show the performance of ensemble model, we used visualization tool, called Receiver operating characteristic (ROC) and Confusion Matrix, which is typically used for model evaluation in the field of artificial intelligence. The area under the ROC Curve is based on a non-parametric statistical sign test, the purpose of which is to estimate the probability of survival for each of a pair of patients. In this study, the area under the ROC was assessed through stratified 10-fold CV-bagging. Each row of the Confusion matrix represents the instances in a predicted class, while each column represents the instances in an actual class. The detailed accuracy of individual models and of the ensemble model is shown in Figure 3.

4 Results and Discussion

In open repair, these patients are likely to have a relatively high postoperative morbidity rate with complications, and will highly influence longer-term postoperative out-comes. It is essential to create reliable and satisfactory risk prediction models for postoperative morbidity as an aid to clinical decision-making. Although several risk prediction systems have been proposed for patients undergoing open aneurysm repair, they basically rely on traditional statistical methods. We have proposed an ensemble model to predict postoperative morbidity after Open Repair and support clinical decision-making. The proposed ensemble model is constructed by Decision Tree, Artificial Neural Network and Support Vector Machine were used to augment the ensemble model and design a Web-Based Application, showing moderate performance.

The experimental result shows our designed Intelligent Analysis Tool can predict and classify the evolution of patients and facilitates different types of users to access functionalities in Weka, the tool can install and access by any client in the network. The proposed ensemble model predicts postoperative morbidity with relatively satisfactory accuracy, even when data is missing and/or sparse, showing its usefulness in support of clinical decision-making. The supplementary nature of multi-models distinguishes the proposed model from existing risk scoring systems that are based on conventional statistical methods and from various machine learning models. To summarize, the advantage of using the proposed ensemble model is that it can provide surgeons with practical, relatively accurate aid in their daily diagnostic tasks.

References

1. Bohm, N., Wales, L., Dunckley, M., Morgan, R., Loftus, I., Thompson, M.: Objective risk-scoring systems for repair of abdominal aortic aneurysms: applicability in Endovascular repair. European Journal of Vascular and Endovascular Surgery 36, 172–177 (2008)
2. Stijn, C.W., Wouters, L.N., Freek, W.A.V., Rene, M.H.J.B.: Preoperative prediction of early mortality and morbidity in coronary bypass surgery. Cardiovascular Surgery 10, 500–505 (2002)

3. Barnes, M., Boult, M., Maddern, G., Fitridge, R.: A model to predict outcomes for Endovascular aneurysm repair using preoperative variables. European Journal of Vascular and Endovascular Surgery 35, 571–579 (2008)
4. Delen, D., Sharda, R., Kumar, P.: Movie forecast Guru: A Web-based DSS for Hollywood managers. Decision Support Systems 43(4), 1151–1570 (2007)
5. Murphy, C.K.: Identifying diagnostic errors with induced decision trees. Medical Decision Making 21(5), 368–375 (2001)
6. Eom, J.-H., Kim, S.-C., Zhang, B.-T.: AptaCDSS-E: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. Expert Systems with Applications 34, 2465–2479 (2008)
7. Polat, K., Güne, S.: Breast cancer diagnosis using least square support vector machine. Digital Signal Processing 17, 694–701 (2007)
8. Babaoğlu, I., Fındık, O., Bayrak, M.: Effects of principle component analysis on assessment of coronary artery diseases using support vector machine. Expert Systems with Applications 37(3), 2182–2185 (2010)
9. Choi, S.: Detection of valvular heart disorders using wavelet packet decomposition and support vector machine. Expert Systems with Applications 35, 1679–1687 (2008)
10. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, San Francisco (2011)

Developing Marketing Strategies Based on Taste Analysis of Mineral Water

Le Yu, Junzo Watada, and Munenori Shibata

Abstract. This research concerns with the development of marketing strategy of mineral water based on people's taste preference by analyzing taste components of mineral water. A two-dimensional analysis has been used in classifying tastes' data. The characteristics of data are recognized in tastes of mineral water by correlation analysis. A combination of Principal Component Analysis and Self-organizing Map is applied to classify the tastes of mineral water. Some marketing strategies are concluded after the evaluation.

Keywords: Taste analysis, mineral water, soft computing model, SOM, Kansei Engineering.

1 Introduction

Kansei engineering, affection engineering or emotion engineering in food and taste field is one branch of the basic sciences in Consumer-oriented Engineering and Production applying specific methods on concrete cases.

This affection engineering field contains two aspects of functions [12]. On one hand, it has been said effective in developing new markets, food industries and even new social systems by applying this advanced technology, which has been put into use in the production and marketing strategies. On the other hand, when understanding consumers' preferences of tastes and smells which are measured by quantitative measurements using special machines, this methodology has been proved helpful in improving new products and making the optimal production processes.

Le Yu · Junzo Watada
Graduate School of IPS, Waseda University, 2-7hibikino, Wakamatsu, Kitakyuusyuu-shi, Fukuoka 808-0135, Japan

Munenori Shibata
Taste & Aroma Strategic Research Institute, Shinkawa Chuou Bldg 8F, 1-17-24, Shinkawa, Chuouku, Tokyo 104-0033 Japan

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 497–507.
springerlink.com

Nowadays, with the development of new technologies, artificial lipid membranes have been invented to imitate biological mechanisms and functions of how human acknowledge taste. According to this invention, now a boom of evaluation methods has come to quantitatively evaluate food tastes in the view of engineering, as well as qualitatively assessments.

Mineral water is water from a mineral spring containing various minerals such as salts and sulfur compounds. Traditionally, mineral water botles were consumed at their spring sources, with the function of "cure". This kind of water is first produced for the underground water sometimes tastes awful or even not fit to drink. Since World War Two, mineral water markets have been successfully created in Western Europe and USA. France leads the world in consumption of mineral waters, and remains the leading producer [8].

For Japanese markets, the consumption of mineral water has been increasing on record. The production quantity inside Japan and imported from oversea both show a growing tendency, in which the oversea import volume occupies an around 20 percent of the whole market till 2011[14]. However, although the growing tendency sustains, consumption per head is still lower than Western countries, such as USA and Canada. So we can see that there is still remaining a huge potential market in Japan. Hence how to produce a bestselling mineral water product could be of great significance. Of course a bestselling product needs a high popularity, but tasty will remain the first important element in our humble opinion.

2 Background and Research Purpose

Generally speaking, there are two branches which have giant differences in analyzing food's tastes, sensory evaluation by human beings' sensory and scientific assessment applied by machine.

Sensory evaluation by human beings' feelings and tastes can be said to be such a subjective evaluation in analyzing food's tastes. This method is based on people's subjective experiences or feelings at the right time, limited by human beings' acknowledges. So it can be very frangible according different people. Even though the same person could have devise tastes if they eat the same food at each time because of different physical conditions, emotions, places and so on. It is said to be very difficult in evaluating tastes in a reliable reproduction.

On the other hand, thanks to the innovation of technique, recently, there has been some special machines can be helpful in improving these mentioned problems, which can be concluded as scientific assessments, such as brix sensor, GC-MS (Gas chromatography - mass spectrometry), HPLC (High Performance Liquid Chromatography), viscoelasticity measuring instruments, etc.

Karterien Beullens, et al. have applied the electronic tongue and attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) for rapid detection of sugars and acids in tomatoes, 2006. Some unsupervised and supervised multivariate data analysis techniques have been applied, such as Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA) [3]. And they compared a new type of electronic tongue in the other paper, also use PCA

method in analyzing tastes of tomatoes [4]. Wei He et al. proved that an electronic tongue is capable of identify teas from different geographical origins and quality grades [6].

Even though we have got these great breakthroughs of taste analyses, there has been a very nonnegligible shortage of mechanical evaluation. The special machines can measure some tastes and smells according to mass, but they cannot measure the quantity of feelings and smells, at which aspect a limitation can be concluded.

In order to solve the problem mentioned above, a new invented taste sensor has been put into use. This taste sensor, which employed lipid polymer membrane, measuring the quantities of human beings' tastes, has been invented recently. And this sensor makes it possible to evaluate food's tastes applied Waive Rule to changing the feelings into quantity values, so that feelings of human beings for food can be said measurable and evaluable.

3 Measurements of Taste

3.1 Taste Sensor and Tastes

Taste sensor can solve the problem that the former instruments could not assess the quantities of feelings. It utilizes a biomimetic membrane, in which lipids are immobilized to the polymer, as a transducer. Then ionization changes of the lipid membrane are detected by electric signals. Following this procedure, the tastes can be recognized after repeating this progress in multi-channel way. Not only the quantities of one fixed taste in the food can be abstracted out by the taste sensor, but the information of taste's weight and qualities can also be acquired.

The flavors can be divided into two groups, fore-taste and after taste, according the time periods that people can feel and recognize it. Fore-taste shows the taste at the moment people eat the food. In another words, it is the sense that the food or beverage is still in the mouth. The after taste shows the continuous feeling when the taste substances are left on the surface of the tongue after the food is swallowed. It is a kind of spread of flavors in one's mouth.

Moreover, there are two attributes included into this research. They are pH value and Electric conductivity. pH value demonstrates the degree of acid and alkali. Electric conductivity illustrates the relations between the total ionic substance foods. The change can be gotten in hand based on this index.

3.2 Calculation Method of Speculation

The calculation is to speculate the different degree of taste based on the putout applied taste sensor. A proportional method is applied to present the strength of discrimination threshold when facing the same kind of stimulus. It is one kind of empirical rules named Waive Rule.

4 Data Analyses of Mineral Water's Taste

In this section, several analyses dealing with mineral water data set are being accomplished. The purpose of these analyses is to figure out some classifications of both the attributes of tastes and the groups of mineral water commodities.

The data set we choose contains 15 dimensions of attributes which present the tastes of mineral water, with 185 samples of different mineral water commodities sold in Japan. These samples' producing area covers all over Japan, and even includes some foreign countries and areas, such as USA and France.

4.1 Correlation Analysis

Correlation analysis is a kind of analyses, and measures the closeness of relationship between two or more variables. In statistics, dependence refers to any statistical relationship between two random variables or two sets of data. Correlation refers to any of a broad class of statistical relationships involving dependence. It is represented by a value within the range of -1.00 to +1.00.

In this case of correlation analysis, we wanted to try to find some highly related attributes so that we can reduce some dimensions.

Full-bodied flavor shows a highest correlation with bitterness as the value of 0.788. This highly correlation proves the common sense that why the sea fish tastes delicious, even though in mineral water's taste.

Taste of Acid flavor A is positively correlated with Miscellaneous bitter taste in medicine, but negatively correlated with Sweetness. The same, Taste of astringency shows a relatively high positive correlation with astringency stimulus and Bitterness in food which is one sort of after tastes. What's more, Salty taste displays a higher correlation with Electric conductivity.

However, the dimension reduction cannot be completed because that the correlation between two attributes is not high enough. Hence, the Principal Component Analysis is being introduced.

4.2 Principal Component Analysis

Principal component analysis (PCA) [1] is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. As a contrast, there are neither any regression analyses nor any objective variables, only compressions of described variables.

Principal component scores are employed to illustrate the characters of each sample and make scatter diagram to show. We set the variance of each principal component is standardized 1.0, for it is easy to make the relationship between variables and samples clearer. Also, in this way, the standardized graph will figure out the relative position of each plot as well as the relationship of variables and samples.

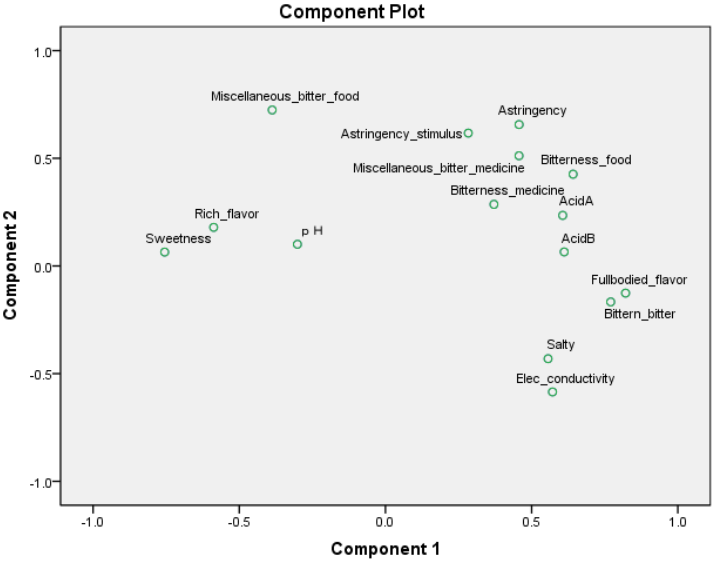


Fig. 1 PCA plot

According to the analysis result, the first component explained about 32% of the variance in the data set of mineral water's tastes, which includes the following elements: bittern bitterness, bitterness in food, full-bodied flavor, and sweetness. So we named this principal component as "fat flavor with bitterness and the-less-the-better sweetness". Mostly these attributes of taste belong to the after taste.

The same as the first component, the second component explained about 17% of the variance, which contains these tastes: miscellaneous bitter taste in food, astringency stimulus and astringency. As the main tastes are stringency and bitterness, we denoted it "astringency and bitterness".

The third principal component is named "salty and pharmic bitter" because of the higher scores of miscellaneous bitter taste in medicine and salty taste.

Figure 1 is composed by the first component and the second components analyzed by PCA method. We can easily figure out that the bitter taste, astringency, and acid taste are pointed in the first quartile. That is to say that most mineral water's tastes are contained of those tastes which can lead a feeling of stimulus. And the third quartile contains of 2 pairs of highly correlated attributes. They lay on this quartile mainly because of the commodity samples we got are mostly manufactured in Japan which cannot get rid of the connection with sea water.

However, from the past researches, we can find out that the PCA method is widely used in analyzing high-dimensional data. But the data volume of these related papers was not very large [6], [7]. Experimental evidence shows that this PCA method can easily divided small amount of data, but the measurement of large dataset is below expectation.

4.3 SOM (Self-Organizing Map)

The SOM algorithm has been invented by Kohonen in 1982 [9], which is a non-supervised neural network and has only two layers, the input layer and the output layer. The input layer is one-dimensional, on the contrary, the output layer is usually two-dimensional and often arranged in away that each unit or neuron is the neighbor of another six, that is why they are represented as hexagons. A neighborhood arrangement is maintained in the SOM algorithm where nearby map neurons have similar profiles after training. Every neuron has a parametric reference vector or prototype vector and the whole input data set in the form of input vectors is presented for training according to the batch version of Kohonen's SOM training algorithm and each observation is projected onto a winning node. The neuron with the prototype vector having the shortest distance from the presented vector is the winning node. The vectors of the winner and its neighboring nodes are modified following the training to represent the input signals in a topology preserving fashion. Finally, the SOM net can project data from an n-dimensional space to a usually two-dimensional grid of neurons and thus facilitate a better visualization of the data [2], [5].

5 Market Analysis and Marketing Strategy

Our research aims at finding out some key elements of tastes that effect marketing profits. To be specific, we focus on figuring out what elements in taste data of dressings will be highly evaluated and drive to high benefit.

5.1 Marketing Analysis Based on PCA

5.1.1 Strategy by the First and Second Component

As Figure 2 shows, the mineral water samples have the tendency of concentration at the second quartile, which represents the first taste into mouth of sweet and stimulus. This result indicates the significance of first taste which shows the moment taste when drinking some mineral water. Companies who manufacture mineral water may change the recipe to make the first taste attractive.

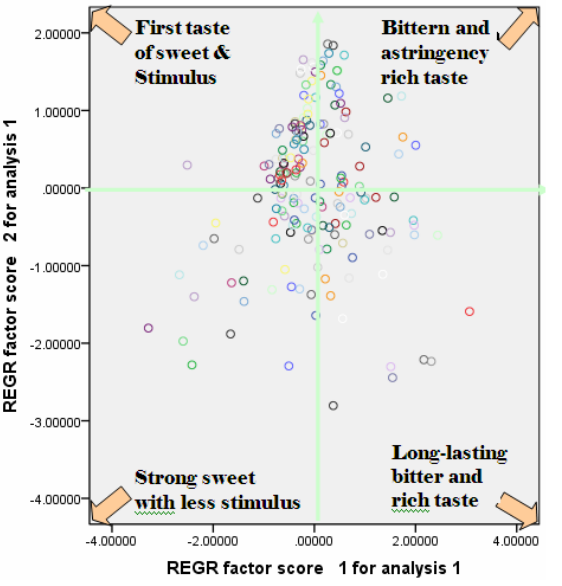


Fig. 2 The 1st and 2nd component distribution plot

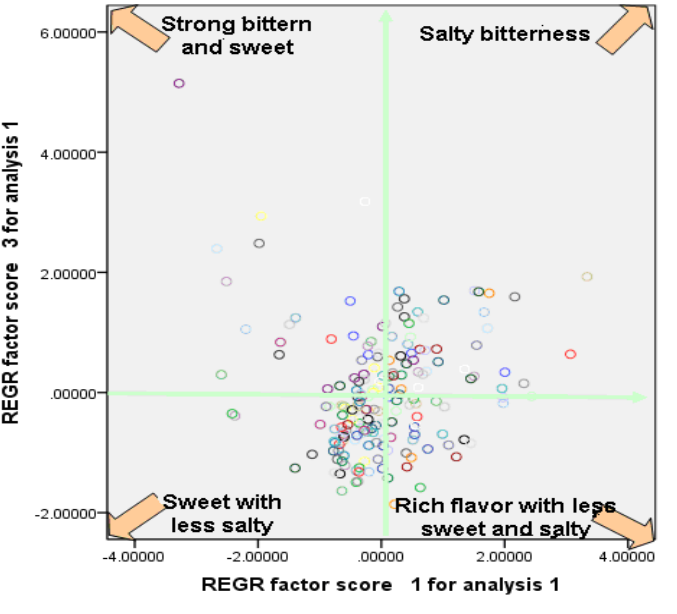


Fig. 3 The 1st and 3rd component distribution plot

5.1.2 Strategy by the First and Third Component

In Figure 3, data points has distributed around the origin point, not that distribute as Figure 2 which described the first and second components. So the importance of illustration is weaker than that one in Figure 2. But in this figure, negative Y-axis axle notably, demonstrates a relatively concentrated tendency. This tendency illustrates that the most mineral water samples have a less salty flavor with weak bitter taste. And moderate sweet flavor is acceptable, for the forth quartile's points.

5.1.3 Strategy by the Second and Third Component

Long-lasting astringency is being preferred in about more than half kinds of samples. People who drink water, especially mineral water, are watching for quenching one's thirst rather than for fun or pursuing sweetness. The feeling of relieving one's throat is being esteemed.

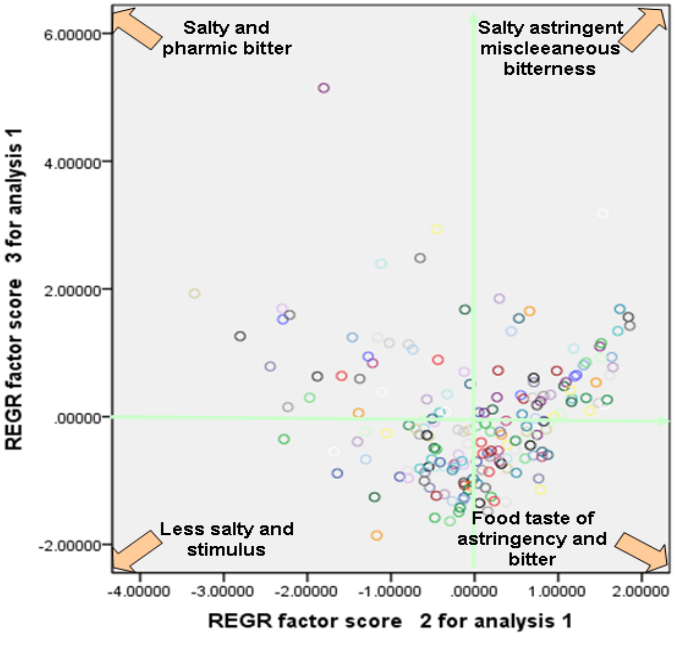


Fig. 4 The 2nd and 3rd component distribution plot

5.2 Marketing Analysis Based on SOM

U-matrix (unified distance matrix) [10] representation of the Self-Organizing Map visualizes the distances between the neurons. The distance between the adjacent neurons is calculated and presented with different colorings between the adjacent

nodes. A dark coloring between the neurons corresponds to a large distance and thus a gap between the codebook values in the input space. A light coloring between the neurons signifies that the codebook vectors are close to each other in the input space. Light areas can be thought as clusters and dark areas as cluster separators. This can be a helpful presentation when one tries to find clusters in the input data without having any a priori information about the clusters.

From Figure 5, we can find out that most of the U-Matrix area is filled by dark blue, see the left-up corner's figure. It shows that a majority of mineral water samples' components are quite similar. Only a minority of these samples present a different category. Hence it can be a little bit difficult in analyzing and classifying the tastes of mineral water.

Among these small pictures, we can see that there are two of them show an obvious variation. They represent taste of Miscellaneous bitter taste in food and Astringency. Therefore, there can be an approach that the manufactures may change the volume of these tastes in order to make their commodity unique, of course in the reasonable scope.

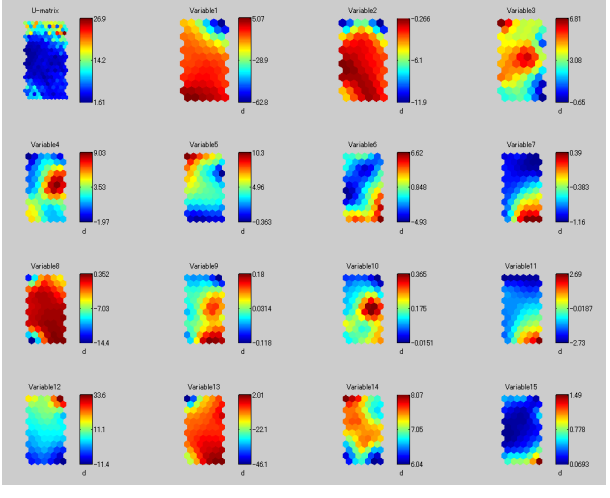


Fig. 5 The U-Matrix of variables

The volume of these tastes in order to make their commodity unique, of course in the reasonable scope.

5.3 Combination of SOM and PCA

With the intention of understanding the relationship of taste and the evaluation, we employed the combination analysis of SOM and clustering [11]. Here we divided the set of taste data into 4 groups after clustering calculation. SOM has been realized by Matlab 7.0.4 with the following parameters.

Learning times: 10000; learning coefficient: 0.01; neighborhood radius: 30; srAnd value: 30; threshold: 0.03

The calculation and operation result figure out about 20 samples among the total 185 samples. And the other kinds of samples can be concluded into one category. These 20 distinctive samples can be also divided into three groups. Samples of NO.29, 36, 90, 95 are proved to be in one category after calculating their distance between each other. From the attributes of these samples, we can easily find out that they are manufactured at the same prefecture, that is, Kagoshima Prefecture in Japan, no matter the hardness of mineral water, nor the depth of collected water. Four samples are gathered into one class, we figured out that these samples are produced out of Japan. And other 10 samples are extracted from seawater, which with higher value of Miscellaneous bitter taste in medicine and Acid flavor A.

According to this classification, we have drawn some conclusions. First of all, mineral water's taste is not decided by producing area. Even though mineral water extracted from the same mountain in the same prefecture, the taste may be different on the basis of taste attribute's values. Secondly, hardness of mineral water does not show influences of taste according to the analysis results.

6 Conclusions

In summary, people who drink mineral water are not for some kinds of excitement, nor for fun. They look for solving the problem of throat's thirst. Hence, the main attributes of mineral water are different from the other beverages, such as fruit drink and soft drink.

This research has concerned about the marketing construction and strategy based on mineral water tastes data set using Correlation, PCA and SOM methods. Section 1 has illustrated the purpose of Kansei Engineering in food area and methods have been applied in analyzing food. It has narrated the proposition that is to apply the taste sensor employing lipid polymer membrane to solve the problem which has been seen irresolvable before in Section 2. Section 3 has demonstrated the characteristics of tastes and taste sensor. Marketing strategies based on mineral water's data character has been discussed in Section 4. Some significant attributes have been pointed out by several methods. Section 5 has described the marketing analyses based on the taste characteristics of mineral water. Some algorithms, such as SOM, PCA, and Correlation have been applied to analysis the nature of tastes in mineral water and try to find some common grounds.

References

[1] Abdi, H., Williams, L.J.: Principal Component Analysis, Wile Interdisciplinary Reviews. Computational Statistics 2(4), 433–459 (2010)

[2] Astel, A., Tsakovski, S., Barbieri, P., Simeonov, V.: Comparison of Self-organizing Maps Classification Approach with Cluster and Principal Analysis for Large Environmental Data Sets. Water Research 41(19), 4566–4578 (2007)

[3] Beullens, K., Kirsanov, D., Irudayaraj, J., Rudnitskaya, A., Legin, A., Nicolai, B.M., Lammertyn, J.: The electronic tongue and ATR-FTIR for rapid detection of sugers and acids in tomatoes. Sensors and Actuators, 107–115 (2006)
[4] Beullens, K., Meszaros, P., Vermeir, S., Kirsanov, D., Legin, A., Buysens, S., Cap, N., Nicolaı, B.M., Lammertyn, J.: Analysis of tomato taste using two types of electronic tongues. Sensors and Actuators, 10–17 (2008)
[5] Giraudel, J.L., Lek, S.: A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. Ecological Modeling 146(1-3), 329–339 (2001)
[6] He, W., Hu, X., Zhao, L., Liao, X., Zhang, Y., Zhang, M., Wu, J.: Evaluation of Chinese tea by the electronic tongue: Correlation with sensory properties and classification according to geographical origin and grade level. Food Research International, 1462–1467 (2009)
[7] Kara, D.: Evaluation of trace metal concentrations in some herbs and herbal teas by principal component analysis. Food Chemistry, 347–354 (2009)
[8] Knox, S., de Chernatory, L.: The application of multi-attribute modeling techniques to the mineral water market. The Quarterly Review of Marketing, School Working Paper SWP 35/ 89
[9] Kohonen, T.: Self-organized Formation of Topologically Correct Feature Maps. Biological Cybemetics 44, 135–140 (1982)
[10] Ultsch, A.: U*-Matrix: a Tool to visualize Clusters in high dimensional Data. Technical Report 36, CS Department, Philipps-University Marburg, Germany (2004)
[11] Vesanto, J., Alhoniemi, E.: Clustering of the Self-organizing Map. IEEE Transactions on Neural Networks 11(3), 586–600 (2000)
[12] Watada, J., Le, Y., Ogura, M., Shibata, M., Fukuda, T.: Building the Marketing Strategies Based on Kansei of Tastes. In: Proceedings of Kansei Engineering Conference at Tokyo, September 3-5 (2011) (in Japanese)
[13] Taste & Aroma Strategic Research Institute (December 1, 2011), http://www.mikaku.jp
[14] The Mineral Water Association of Japan (January 26, 2012), http://minekyo.net/index.php

Estimation of Subjective Stress via Simplified Biosignal Measurement

Yoshihito Maki, Genma Sano, Yusuke Kobashi, Tsuyoshi Nakamura, Masayoshi Kanoh, and Koji Yamada

Abstract. We have studied an estimation of subjective stress using the finger plethysmography so far. However the measurement of the finger plethysmography makes the subjects feel suffering or physical load. This paper proposes a method via NeuroSky MindSet headset to reduce the suffering or the physical load, instead of the method using the finger plethysmography. We conducted the experiment to compare the estimations of the subjective stress via both the MindSet and the sensor of finger plethysmography. The experimental results indicated that the biosignals via MindSet could have higher potential than the finger plethysmography to estimate the subjective stress.

1 Introduction

In these days the interdisciplinary field of affective computing, the study and development of systems and devices that can recognize, interpret, and simulate the emotional states of human, has attracted enormous interest [1, 2]. The range of data available for extrapolating human emotion includes not only brain-wave or electrocardiogram (EEG) patterns, electrocardiographic (ECG) patterns, and electromyogram (EMG) patterns, but also the voice, facial expressions, posture, and behavior. In choosing which type of data to observe, facial expression and speech have obvious advantages since this data can be obtained by contactless sensors, and thus

Yoshihito Maki · Genma Sano · Yusuke Kobashi · Tsuyoshi Nakamura
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi, Japan
e-mail: {maki,sano,kobashi,tnaka}@ai.nitech.ac.jp

Masayoshi Kanoh
Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota, Aichi, Japan
e-mail: mkanoh@sist.chukyo-u.ac.jp

Koji Yamada
Institute of Advanced Media Arts and Sciences, 3-95 Ryoke-cho, Ogaki, Gifu, Japan
e-mail: k-yamada@iamas.ac.jp

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 509–517.
springerlink.com

imposes little burden on the user. But to achieve greater accuracy, we would expect data robustness to be enhanced by using EEG or EMG sensors that are in contact with the body. Among the various types of data and methods that are available, we wanted the accuracy that can be obtained with contact sensors yet we did not want to be overly dependent on the measurement environment, so we focused on brain-wave (EEG) biosignals. Besides brain-wave signals that we use in this study, there are various other kinds of biosignals captured by contact sensors-electrocardiogram (ECG), electromyogram (EMG), electrodermogram (EDG), and so on-but brain waves directly reflect activities of the brain, which should also include information relating to the state of human emotion.

Measuring brain waves in the past was a cumbersome procedure. It involved an enormous expensive piece of equipment and numerous electrodes were attached to the subject's skull with copious amounts of gel. While this approach is still in widespread use, we've recently seen the emergence of a much simpler brain-wave sensor that uses dry (i.e., no gel) electrodes. These simpler sensors support mobile applications, and if they can be further downsized, one can easily imagine a wearable implementation that could be conveniently worn throughout the day.

We propose the use of a biological signal (biosignal) sensor for measuring and estimating stress. Specifically, we envision a mechanism through which a robot observes the signals output from a biosignal sensor worn by a human being. There are a variety of indices for estimating stress, but we here explore the feasibility of estimating stress based on the biosignals detected by the MindSet headset from NeuroSky, Inc.[3, 4, 5, 6]. We examine, in particular, the relationship between perceivable subjective stress and the biosignals obtained from MindSet.

We point out here that long-term or unperceivable stress is outside the scope of this study that targets moderate and short-term stress that occurs in everyday life. In this regard, we readily acknowledge that long-term or unperceivable stress can be extremely meaningful in the research of mechanisms for detecting danger and avoiding accidents. However, such non-moderate stress that occurs over the medium or long term is more applicable to making medical diagnoses and providing medical treatment; it is a condition that should be dealt with by a medical specialist rather than a robot.

2 Estimation of Subjective Stress

Substances present in saliva such as the immunoglobulin-A antibody and cortisol, a hormone, are widely known as indices for making physiological evaluations of stress [7, 8, 9]. It is also known that stress can be evaluated by heart-rate variability analysis. Yokoyama et al. [10] have proposed a method for estimating the high-frequency component of a heart rate called the turning point amplitude (TPA), which is a kind of heart-rate variability analyses and recognized as having a high correlation with the conventional auto regressive (AR) method and analysis by wavelet transformation. The TPA can be calculated from 20 - 30 points of data and can be applied to real-time processing thanks to a relatively simple formula.

In past research, we investigated simple non-invasive techniques for measuring subjective stress at a level that could be sensed by humans and explored the application of those techniques [11]. To evaluate subjective stress, we calculated, in particular, TPA from finger plethysmography (TPA_f), which is an easier process than measuring heart rate. On examining the relationship between TPA_f and subjective stress, we confirmed that the presence or absence of subjective stress could be judged on the basis of an objective index. This method, however, is not considered to be very practical since the subject has to wear a pulse-wave sensor on a finger and must keep that hand stationary to measure stress during plethysmography. It is therefore thought to be limited with regard to future applications. Another problem that remains is that the ability of this method to estimate subjective stress "to some extent" is simply not sufficient for our research objectives. In the study presented here, we address these problems by proposing a technique that uses the MindSet device from NeuroSky. As shown in Fig. 1, MindSet is worn like a headphone, which eliminates any restrictions in hand movement. The user therefore has a higher degree of freedom with MindSet compared to a finger-worn pulse-wave sensor. Furthermore, compared to standard electroencephalography (EEG) equipment that attaches many wet electrodes to the subject for measuring brain waves, MindSet requires the use of only one dry electrode. This and the fact that it can be used wirelessly make MindSet especially applicable to mobile applications, and we envision its use in everyday scenarios. Despite having such a simple set of specifications, the MindSet format has been shown to be effective as a steady-state visual evoked potential (SSVEP)-based brain-computer interface (BCI) [5], which suggests that MindSet has the potential of being used for estimating subjective stress. In this paper, we explore by experiment the feasibility of estimating subjective stress using biosignals obtained from MindSet. Our findings suggest the possibility that MindSet can be used to estimate subjective stress with greater accuracy than the use of TPA_f.

3 Subjective-stress Estimation Experiment

We conducted the following experiment to compare the use of finger plethysmography and MindSet in estimating subjective stress.

3.1 Experimental Setup and Procedure

The experiment was conducted in a quiet, closed-off room with 16 male subjects in their 20s. These subjects were briefed on the nature of the experiment beforehand and were given a chance to practice mental arithmetic (the stress-load task in this experiment) for one minute. Each subject was seated during the experiment wearing a MindSet headset with a finger plethysmography device attached to the left hand.

The experimental time chart is shown in Fig. 2. In the experiment, a five-minute rest state (R_i) with eyes closed and a five-minute stress state (S_i) performing mental arithmetic were taken to be a set of two phases. This set is repeated five times in a row ($i = 1, \ldots, 5$) so that 50 minutes were required to perform this experiment

once. In asking subjects to perform mental arithmetic five times, we referred to the study presented in "The Variations of Tenseness and Heart Rate during the Repeated Mental Arithmetic Tasks" [12]. The mental arithmetic performed in the stress state consisted of counting down from a four-digit integer displayed on the screen of a personal computer in increments of 13 using a mouse for inputting answers (Fig. 3). The initially displayed four-digit integer is randomly selected, and input of an erroneous answer causes the system to return to and redisplay the original four-digit integer. This mental arithmetic system is based on the "Trier Social Stress Test (TSST)" [13].

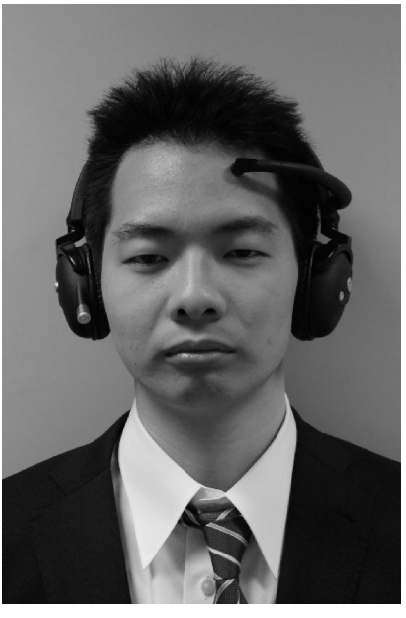


Fig. 1 Neurosky MindSet

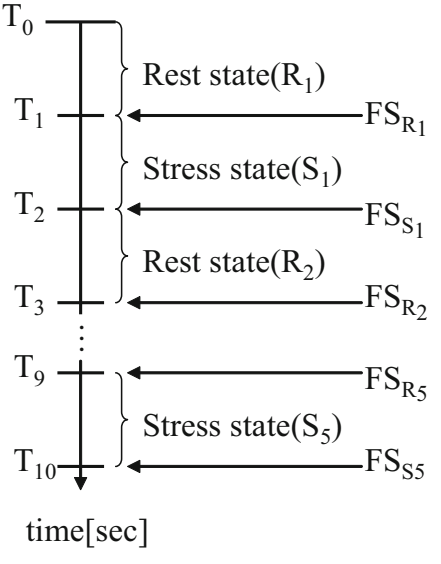


Fig. 2 Experimental time chart

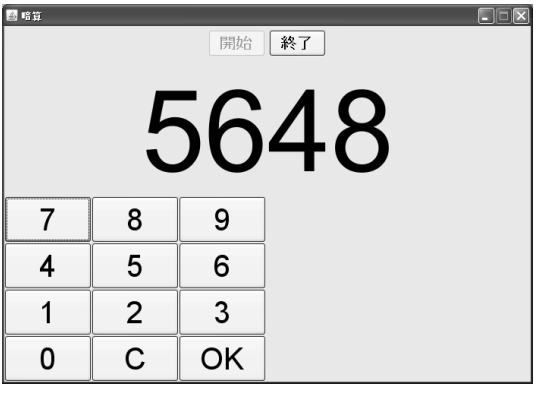


Fig. 3 Mental arithmetic application

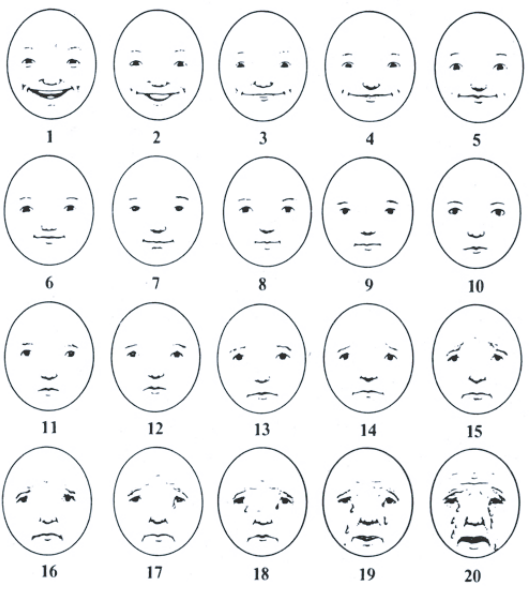


Fig. 4 Face Scale

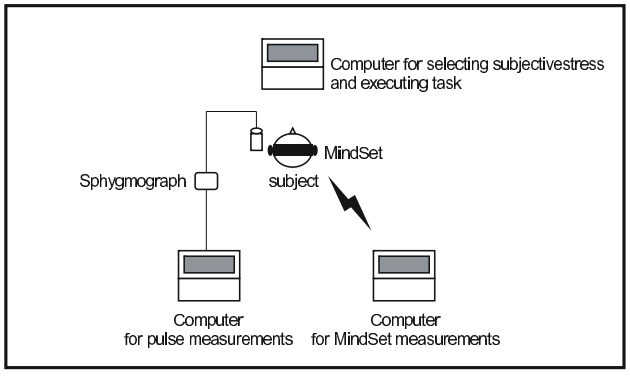


Fig. 5 Experimental setup

Additionally, at the conclusion of each state (at time points $T_1, \ldots, T_{10}$, the psychological state of the subject was subjectively evaluated using the face-scale method developed by Lorish et al. [14] and the result was taken to be the subjective stress of the subject at that time. The face scale is a subjective assessment scheme consisting of 20 drawings of a single face, with each face depicting a slightly different mood state numbered from 1 to 20, with 1 representing the most positive mood (a very happy face) and 20 representing the most negative mood (a very sad face). Subjects can very quickly and easily subjectively assess the way they are feeling

throughout the experiment by simply indicating the face that best reflects how they are feeling at the moment. The subjects were instructed to choose the face from the face scale that best represented their mood or how they were feeling at the moment. Figure 4 shows the face scale that was used in this work.

Biosignals from both finger plethysmography and MindSet were also obtained at each time point. The experiment was carried out using the setup shown in Fig. 5 and was performed once for each subject according to the procedure shown in Fig. 2.

3.2 Stress Evaluation Method

This study evaluated the time series of data obtained for each biosignal in the following way. First, we obtained eight types of biosignals $\mathbf{M_8} = (\delta, \theta, \alpha_h, \alpha_l, \beta_h, \beta_l, \gamma_h, \gamma_l)$ which indicates the frequency spectrum pattern of the biosignals from MindSet once every second, and we calculated $\mathrm{TPA_f}$ as the average value of a time series of pulse-wave variability over a fixed period of time. Fig. 6 shows an example of time series of $\mathbf{M_8}$ and $\mathrm{TPA_f}$.

We considered that a time interval directly before the end of each state shown in Fig. 2 to be the most appropriate time period for explaining that state. Accordingly, for the initial rest state $\mathrm{R_1}$, we defined the biosignals obtained over τ seconds directly before end time $\mathrm{T_1}$ as $\mathrm{X_{R_1(\tau)}}$, and for the initial stress state $\mathrm{S_1}$, we defined the biosignals obtained over τ seconds directly before end time $\mathrm{T_2}$ as $\mathrm{X_{S_1(\tau)}}$.

In general, then, the biosignals corresponding to the n-th rest state $\mathrm{R_n}$ were defined as $\mathrm{X_{R_n(\tau)}}$ and the biosignals corresponding to the m-th stress state $\mathrm{S_m}$ were defined as $\mathrm{X_{S_m(\tau)}}$. In this experiment, we set τ to 30 s. Here, we did not use the absolute values of biosignals considering differences in the psychological states of subjects at the time of the experiment. Instead, treating the first rest state $\mathrm{R_1}$ as the base rest state, we used the difference between the biosignal in question with the average value of $\mathrm{X_{R_1(\tau)}}$, that is, $\overline{\mathrm{X}}_{\mathrm{R_1(\tau)}}$, to compute a variation, and for each state, we defined average variation as , $\Delta\overline{\mathrm{X}}_{\mathrm{R_n(\tau)}}\mathrm{C}\ \Delta\overline{\mathrm{X}}_{\mathrm{S_m(\tau)}}$ as given by Eqs. (1) and (2). We performed our evaluation based on these values.

$$\Delta\overline{\mathrm{X}}_{\mathrm{R_n(\tau)}} = \frac{1}{\tau}\sum_{\mathrm{t=T_{2n-1}-\tau}}^{\mathrm{T_{2n-1}}}\{\mathrm{X_{R_n(\tau)}(t)} - \overline{\mathrm{X}}_{\mathrm{R_1(\tau)}}\} \tag{1}$$

$$\Delta\overline{\mathrm{X}}_{\mathrm{S_m(\tau)}} = \frac{1}{\tau}\sum_{\mathrm{t=T_{2m}-\tau}}^{\mathrm{T_{2m}}}\{\mathrm{X_{S_m(\tau)}(t)} - \overline{\mathrm{X}}_{\mathrm{R_1(\tau)}}\} \tag{2}$$

$$n \geq 2, m \geq 1, \mathrm{X} \in \{\delta, \theta, \alpha_h, \alpha_l, \beta_h, \beta_l, \gamma_h, \gamma_l, \mathrm{TPA_f}\}$$

Values obtained by the face-scale method were defined as follows. The face-scale value obtained at end time $\mathrm{T_1}$ of the first rest state was defined as $\mathrm{FS_{R_1}}$ and that obtained at end time $\mathrm{T_2}$ of the first stress-load state was defined as $\mathrm{FS_{S_1}}$. In general, then, the face-scale value after the n-th rest state was defined as $\mathrm{FS_{R_n}}$ and that after the m-th mental-arithmetic state was defined as $\mathrm{FS_{S_m}}$. As in the case of biosignals, we here used the difference values given by Eqs. (3) and (4) in which the face-scale value after the first rest state is used as a base value.

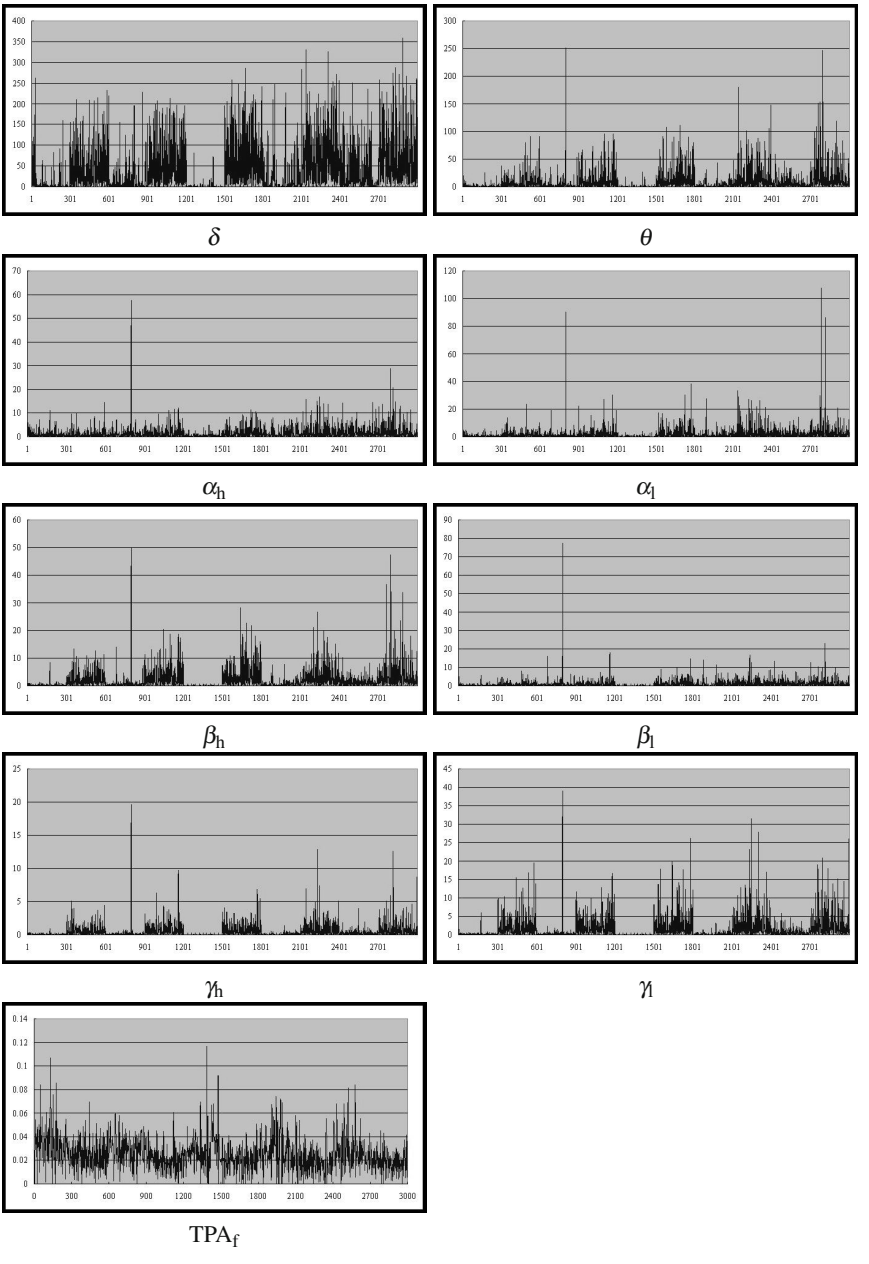


Fig. 6 Time series of $\mathbf{M_8}$ and TPA_f

$$\Delta FS_{R_n} = FS_{R_n} - FS_{R_1} \tag{3}$$

$$\Delta FS_{S_m} = FS_{S_m} - FS_{R_1} \tag{4}$$

4 Experimental Results and Discussion

We performed multiple regression analysis using $\Delta\overline{X}_{R_n(\tau)}$ and $\Delta\overline{X}_{S_m(\tau)}$ as explanatory variables and ΔFS_{R_n} and ΔFS_{S_m} as corresponding objective variables. In our evaluation, we used all variables instead of selecting particular ones. Table 1 lists the coefficient of determination R^2 obtained when performing multiple regression analysis for $\Delta\overline{TPA_f}C\Delta\overline{\mathbf{M_8}}$, and $(\Delta\overline{TPA_f}, \Delta\overline{\mathbf{M_8}})$ each as explanatory variables. Here, $\Delta\overline{\mathbf{M_8}}$ denotes average difference in the eight biosignals obtained from MindSet as expressed by $\Delta\overline{X}_{R_n(\tau)}$ and $\Delta\overline{X}_{S_m(\tau)}$. Examining these results, we see that R^2 when performing a regression analysis on only $\Delta\overline{TPA_f}$ is low at 0.0125, which indicates that estimating subjective stress from only $\Delta\overline{TPA_f}$ would be difficult. In contrast, R^2 when performing a regression analysis on $\Delta\overline{\mathbf{M_8}}$ is 0.2558, which suggest the possibility of estimating subjective stress from $\Delta\overline{\mathbf{M_8}}$ rather than $\Delta\overline{TPA_f}$. These results also show that R^2 for $(\Delta\overline{TPA_f}, \Delta\overline{\mathbf{M_8}})$ is higher than that of $\Delta\overline{\mathbf{M_8}}$, but since the difference there is hardly as large as that between $\Delta\overline{TPA_f}$ and $\Delta\overline{\mathbf{M_8}}$, we consider that there would be little advantage in using both MindSet and a finger pulse-wave sensor together.

Table 1 Coefficient of determination

Explanatory variable	Coefficient of determination R^2
$\Delta\overline{TPA_f}$	0.0125
$\Delta\overline{\mathbf{M_8}}$	0.2558
$(\Delta\overline{TPA_f}, \Delta\overline{\mathbf{M_8}})$	0.2566

5 Conclusion

In this paper, we examined the relationship between biosignals obtained from NeuroSky MindSet headset as an objective evaluation and the results of applying the face scale as a subjective evaluation. The results of this survey show the potential of using these biosignals for estimating subjective stress. Using biosignals from MindSet instead of using TPA_f obtained from the sensor of finger plethysmography as in previous research eases the burden on the subject by making it unnecessary to wear a pulse-wave sensor on a finger and to keep that hand stationary. In future research, we plan to analyze experimental data in more detail with the aim of improving the accuracy of estimation.

References

1. Picard, R.W.: Affective Computing. The MIT Press (1997)
2. Calvo, R.A., D'Mello, S.: Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing 1(1), 18–37 (2010)

3. Rebolledo-Mendez, G., Dunwell, I., Martínez-Mirón, E.A., Vargas-Cerdán, M.D., de Freitas, S., Liarokapis, F., García-Gaona, A.R.: Assessing NeuroSky's Usability to Detect Attention Levels in an Assessment Exercise. In: Jacko, J.A. (ed.) HCI International 2009, Part I. LNCS, vol. 5610, pp. 149–158. Springer, Heidelberg (2009)
4. Yasui, Y.: A Brainwave Signal Measurement and Data Processing Technique for Daily Life Applications. Journal of Physiological Anthropology 28(3), 145–150 (2009)
5. Luo, A., Sullivan, T.J.: A user-friendly SSVEP-based brain computer interface using a time-domain classifier. Journal of Neural Engineering 7(2), 026010 (2010)
6. Haapalainen, E., Kim, S., Forlizzi, J.F., Dey, A.K.: Psycho-physiological measures for assessing cognitive load. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Ubicomp 2010, pp. 301–310 (2010)
7. Tsujita, S., Morimoto, K.: Secretory IgA in saliva can be a useful stress marker. Environmental Health and Preventive Medicine 4, 1–8 (1999)
8. Valdimarsdottir, H.B., Stone, A.A.: Psychosocial factors and secretory immunoglobulin A. Critical Reviews in Oral Biology and Medicine 8(4), 461–474 (1997)
9. Kirschbaum, C., Hellhammer, D.H.: Salivary cortisol in psychoneuroendocrine research: Recemt developments and applications. Psychoneuroendocrinology 19(4), 313–333 (1994)
10. Yokoyama, K., Kamiya, N.: Simplified method for estimating high-frequency component of heart rate variability based on local characteristics in time series. Ergonomics 40(6), 302–308 (2004)
11. Sugiura, T., Okudaira, Y., Kobashi, Y., Nakamura, T., Itoh, H.: Relaxation effects of object speed in video images: Evaluation by heart rate variability analysis. Kansei Engineering and Emotion Research, in CD-ROM (2009)
12. Nagamine, M., Murota, M., Ishii, M., Shimizu, Y.: The Vaiations of Tenseness and Heart Rate during the Repeated Mental Arithmetic Tasks. Japan Journal of Educational Technology 25(4), 237–245 (2002) (in Japanese)
13. Kirschbaum, C., Pirke, K.M., Hellhammer, D.H.: The 'Trier Social Stress Test'- A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. Neuropsychobiology 28, 76–81 (1993)
14. Lorish, C.D., Maisiak, R.: The face scale:a brief, nonverbal method for assessing patient mood. Arthritis Rheum 29, 906–909 (1986)

Evaluating Cloud Computing Based Telecommunications Service Quality Enhancement by Using a New Hybrid MCDM Model

Chi-Yo Huang, Pei-Chu Hsu, and Gwo-Hshiung Tzeng

Abstract. The fast emergence of the cloud based computation fulfills modern telecommunication serve firms' requirements for massive computation and storage space with very low costs. However, how the cloud-based service quality can be enhanced, how customers' needs can be fulfilled, and how firms' profits can be maximized were seldom discussed. Thus, this research aims to provide a new hybrid MCDM model combining DEMATEL and ANP based framework and Grey Relational Analysis (GRA) for improving to close the service quality gap, fulfill customers' satisfactory and maximize profits in interdependence and feedback problems among dimensions/criteria. An empirical case study bases on enhancing the service quality of a Taiwanese mobile telecommunication firm's application of cloud computation as the billing system. In results of this case study, the expertise in dynamic capacity management and new education and training should be the most important strategies for service quality enhancement to be introduced.

Keywords: Cloud computing, cloud-based services, infrastructure as a service (IaaS), service quality, DEMATEL (Decision Making Trail and Evaluation Laboratory), hybrid MCDM (Multiple Criteria Decision-Making), GRA (Grey Relational Analysis).

Chi-Yo Huang · Pei-Chu Hsu
Department of Industrial Education, National Taiwan Normal University
No. 162, Hoping East Road I, Taipei 106, Taiwan
e-mail: cyhuang66@ntnu.edu.tw

Gwo-Hshiung Tzeng
Graduate Institute of Project Management, Kainan University
No. 1, Kainan Road, Luchu, Taoyuan County 338, Taiwan
e-mail: ghtzeng@mail.knu.edu.tw
and
Institute of Management of Technology, National Chiao Tung University Ta-Hsuch Road, Hsinchu 300, Taiwan
e-mail: ghtzeng@cc.nctu.edu.tw

J. Watada et al. (Eds.): Intelligent Decision Technologies, Vol. 1, SIST 15, pp. 519–536.
springerlink.com

1 Introduction

The cloud computing, a new term for a long-held dream of computing as a utility (Parkhill, 1966), has recently emerged as a commercial reality (Armbrust, Fox, Griffith, Joseph, Katz, Konwinski & Lee, 2009). The cloud computing enables the convenient, on-demand network access to a shared pool of the configurable computing resources which can be rapidly provisioned and released with minimal management effort or service provider interaction (Mell & Grance, 2009). The cloud computing implies a service-oriented architecture, reduced information technology overhead for the end-user, greater flexibility, reduced total cost of ownership, on-demand services and many other things (Vouk, 2008).

As the cloud computation emerged, service firms begin to introduce the cloud based computations computing in providing services by increasing the efficiency while reducing the costs and enhancing the convenient. Albeit important, very few scholars tried to uncover how the cloud based service quality can be also enhanced.

In this research, the authors tried to find some possible ways that may improve cloud computing based telecommunications service quality.

Thus, this research aims to define a new hybrid MCDM model combining DEMATEL and ANP based framework and Grey Relational Analysis (GRA) framework for deriving factors influencing the cloud based service quality in interdependence and feedback problems among dimensions/criteria. Further, appropriate strategies for enhancing the service mechanism to achieve the aspiration level will be provided.

An empirical case study based on enhancing the service quality of a Taiwanese mobile telecommunication firm's application of cloud computation as the billing system. Inresults, Scalable storage and Scaling quickly can be found that have the highest influential weight by ANP. They are 0.499 and 0.342. Furthermore, new education and training with the highest score, 0.860, was derived by using the GRA. In general, scalable storage and scaling quickly are the two serious problems in this case, and new education and training is the suitable way to enhance and improve the service quality.

The reminder of this paper is organized as follows. In Section 2 some scientific or technical literature which is related to cloud computing and strategies for enhancing the cloud based service quality will be reviewed. Then the methods which used in the research are introduced in Section 3. In this paper, a hybrid MCDM model combining DEMATEL, ANP and GRA will be used to improve and enhance telecommunication service quality as the research tools. In section 4, the outcome calculating by DEMATEL, ANP and GRA can be showed in real empirical case study. Then the outcome will be discussed in detail in Section 5. Finally, a conclusion can be reached in Section 6.

2 Literature Review

In the following Section, researches being related the cloud computing, the present state of the cloud computing industry will be reviewed. The review results will serve as a basis for the theoretical framework development.

2.1 What Is Cloud Computing?

Cloud computing has been used in the IT (Information Technology) sector and elsewhere like a buzzword but there is no single definition for the concept as such (Myoken, 2009) Cloud computing is associated with a new paradigm for the provision of computing infrastructure. This paradigm shifts the location of this infrastructure to the network to reduce the costs which were associated with the management of both hardware and software resources (Vaquero, Merino, Caceres & Lindner, 2009).

According to some experts, cloud computing is a kind of computing technique where IT services are provided by massive low-cost computing units connected by IP (Information Provider) networks. Cloud computing is rooted in search engine platform design. There are five major technical characteristics of cloud computing: (1) large scale computing resources, (2) high scalability and elasticity, (3) shared resource pool (virtualized and physical resource), (4)dynamic resource scheduling, and (5) general purpose (Vaquero, Merino, Caceres & Lindner, 2009).

According to the type of provided capability, there are three kinds of scenarios that cloud computing can be used:

(1) **Software as a Service (SaaS):** There are services of potential interest to a wide variety of users hosted in Cloud systems. This is an alternative to locally run applications. This scenario is called Software as a Service (SaaS) (Myoken, 2009).
(2) **Platform as a Service (PaaS):** Cloud systems can offer an additional abstraction level, instead of supplying a virtualized infrastructure. They can provide the software platform where systems run on. The sizing of the hardware resources demanded by the execution of the services is made in a transparent manner.
(3) **Infrastructure as a Service (IaaS):** IPs manage a large set of computing resources, such as storing and processing capacity. Through virtualization, they are able to split, assign and dynamically resize these resources to build ad-hoc systems as demanded by customers, the SPs (service providers). They deploy the software stacks that run their services.

Cloud computing is one of the most vague technique terminologies in history. One reason is that cloud computing can be used in many application scenarios. The other reason is that cloud computing is hyped by lots of companies for business promotion (Vaquero, Merino, Caceres & Lindner, 2009) .

2.2 Cloud Computing and Enterprise

Large enterprises inevitably have a large number of computing systems that have been developed over a long period of time (K-Hosseini, Sommerville & Sriram, 2010). The large telecommunications company said that by using the cloud computing, it has been able to reduce capital expenditures for the affected applications by 30% and achieve significant headcount reduction for systems administration. It said it has been able to mix and match and repurpose the cloud infrastructure far more than any previous infrastructure deployment (Staten, 2009).

It seems perfect that when an enterprise use cloud as its running way. It can reduce the satisfy business requirements on demand, lower cost and energy-saving, improve and enhance the efficiency of resource management (Vaquero, Merino, Caceres, & Lindner, 2009). However, there are still some problems when we are talking about using cloud in the enterprise.

2.3 Service Quality

Service quality has become a critical competitive consideration in cloud computing. What we have come to know as service quality relies on a simple but powerful discrepancy paradigm: initial service expectations must be met or exceeded by perceived outcomes of the service experience (Headley & Miller, 1993).

In general, customer expectations of service quality, Parasuraman, Berry & Zeithaml (1991) points out: (1) **reliability**, the ability of perform the promised service dependably and accurately; (2) **tangibles**, the appearance of physical facilities, equipment, personnel, and communication materials; (3) **responsiveness**, the willingness to help customers and provide prompt service; (4) **assurance**, the knowledge and courtesy of employees and their ability to convey trust and confidence; and (5) **empathy**, the caring, individualized attention provided to the customer. But in our empirical real situation as case study, we base on enhancing the service quality of a Taiwanese mobile telecommunication firm of cloud computation as the billing system. In this case service quality includes: (1) privacy and security (C_1), continuity of service (C_2), service migration (C_3), scalable storage (C_4), scaling quickly (C_5), organizational change (C_6). We will adopt these criteria as example for applying to the proposed new hybrid MCDM model.

2.4 Strategies for Enhancing the Cloud Based Service Quality

Now every enterprise use "Infrastructure as a Service (IaaS)" as the way to improve their hardware. When we use IaaS, we would face some obstacles, like privacy and security, the continuity of service, service migration (Vaquero, Merino, Caceres & Lindner, 2009), scalable storage and scaling quickly (Myoken, 2009), which means dynamic resource allocation. If they want to modify any short in the business, they should overcome these kinds of problems to raise the efficiency.

Another big problem that is the organizational change brought about with cloud computing. Traditional management is not suitable for those enterprises because they introduced cloud computing into their business. It is totally a different way to run a business. Yanosky (2008) discussed how cloud computing will affect the authority of the IT department within universities; however, this also applies to IT departments in enterprise. The IT department gained its authority in the early days of computing when they had the majority of the programming skills and control of mainframes within an organization. As the use of IT expanded within organizations, system administrators and developers were forced to learn new skills as their role was no longer just about keeping the technology running. Until the invention of the PC, users relied on the services provided by the IT department for

systems support. The type of organizational change that cloud computing results in can be demonstrated by taking a look at, for example, IT procurement within an enterprise. Simplistically, procurement is based on obtaining estimates for things, then getting those estimates signed-off by management to allow the procurement to proceed. Capital and operational budgets are kept separate in this process, and it can take several months between the decision to procure hardware and the hardware being delivered, setup and ready to use. The use of cloud computing can greatly reduce this time period, but the more significant change relates to the empowerment of users and the diffusion of the IT department's authority as pointed out by Yanosky (2008).

Therefore, a necessary condition for a company to become a Cloud Computing provider is that it must have been existing investments not only in very large data centers, but also in large-scale software infrastructure and operational expertise required to run them. Given these conditions, a variety of factors might influence these companies to become Cloud Computing providers: (1) make a lot of money, (2) leverage existing investment, (3) defend a franchise, (4) leverage customer relationships, (5) become a platform (Armbrust, Fox, Griffith, Joseph, Katz, Konwinski & Lee, 2009).

In order to manage the data center efficiently, vastly superior economics, better practices for handling dynamic workloads, expertise in dynamic capacity management, new education and training may hold the solution (Staten, 2008).

3 Research Method

In this paper, the authors use a new hybrid MCDM model combining DEMATEL, ANP to find the influential weights of every criterion, and then use GRA to find the suitable strategy in real world. Following, the authors are going to introduce the research method.

3.1 The DEMATEL Technique

The basic concept of DEMATEL technique was developed by the Battelle Geneva Institute to analyze complex 'world problems' dealing mainly with interactive man-model techniques (Gabus, & Fontela, 1972); then in recent decade DEMATEL technique was used to evaluate qualitative and factor-linked aspects/criteria of societal network inter-relationship problems (Tzeng, Chiang, & Li, 2007). The applicability of the method is widespread, ranging from industrial planning and decision-making to urban planning and design, regional environmental assessment, analysis of world problems, and so forth. It has also been successfully applied in many situations, such as marketing strategies, control systems, safety problems, developing the competencies of global managers and group decision-making (Liou, Tzeng, & Chang, 2007; Chiu, Chen, Tzeng, & Shyu, 2006; Wu, & Lee, 2007; Lin, & Wu, 2008). Furthermore, a hybrid MCDM model combining DEMATEL and DANP (DEMATEL-based ANP) has been widely used in

various fields, for example, e-learning evaluation (Tzeng, Chiang, & Li, 2007), airline safety measurement (Wu, & Lee, 2007), innovation policy portfolios for Taiwan's SIP Mall (Huang, Shyu, & Tzeng, 2007), and many others applications (Liou & Tzeng, 2011; Liou, Tsai, Lin & Tzeng, 2011; Yang & Tzeng, 2011; Shen, Lin & Tzeng, 2011; Chen, Hsu & Tzeng, 2011; Chen & Tzeng, 2011; Tzeng & Huang, 2012; Liu, Tzeng & Lee, 2012; Liu, Tzeng & Lee, 2012; Wang & Tzeng, 2012; Ou Yang, Shieh & Tzeng, 2012; Liou, Tzeng, Hsu & Liou, Tzeng, Hsu & Yeh, 2012; Teng & Tzeng, 2012; Kuan, Tzeng & Hsiang, 2012). Therefore, in this paper we use DEMATEL not only can be to detect complex relationships and build a NRM (Network Relation Map) of the dimensions/criteria, but also can be to obtain the influential degree of each dimension/criterion over others in cloud computing service quality; we also can then adopt the normalized influence matrix by criteria in each dimension and transpose it into un-weighted supermatrix using basic concept of ANP (Saaty, 1996). We use the normalized influence matrix by dimensions to multiply un-weighted supermatrix for determining DANP to obtain the relatively influential weights. To apply the DEMATEL method smoothly, the authors refined the definitions based on above authors, and produced the essential definitions indicated below. The DEMATEL method is based upon graph theory, enabling us to plan and solve problems visually, so that we may divide multiple criteria into a relationship of cause and effect group, in order to better understand causal relationships. Directed graphs (also called digraphs) are more useful than directionless graphs, because digraphs will demonstrate the directed relationships of sub-systems. A digraph typically represents a communication network, or a domination relationship between individuals, etc. Suppose a system contains a set of elements/criteria, $S = \{s_1, s_2, \ldots, s_n\}$, and particular pair-wise relationships are determined for modeling, with respect to a mathematical relationship (MR). Next, portray the relationship MR as a direct-relation matrix that is indexed equally in both dimensions by elements/criteria from the set *S*. Then, extract the case for which the number appears in the cell (i, j), if the entry is a positive integral that has the meaning of: the ordered pair (s_i, s_j) is in the relationship MR; it has the kind of relationship regarding that element such that s_i causes element s_j. The digraph portrays a contextual relationship between the elements of the system, in which a numeral represents the strength of influence (Figure 2). The elements s_1, s_2, s_3 and s_4 represent the factors that have relationships in Figure 1. The number between factors is influence or influenced degree. For example, an arrow from s_1 to s_2 represents the fact that s_1 influences s_2 and its influenced degree is two. The DEMATEL technique can convert the relationship between the causes and effects of criteria into an intelligible structural model of the system (Chiu, Chen, Tzeng, & Shyu, 2006; Tzeng, Chiang, & Li, 2007) as following steps.

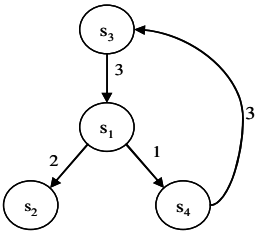


Fig. 1 An example of the directed graph

Step 1: The pair-wise comparison scale may be designated as eleven levels, where the scores 0,1,…,4 represent the range from 'no influence (0)' to 'very high influence (4)'.

Step 2: The initial direct relation/influence matrix $\boldsymbol{A}$ is an $n \times n$ matrix obtained by pair-wise comparisons, in terms of influences and directions between the innovation competences, in which a_{ij} is denoted as the degree to which the i^{th} criteria affects the j^{th} criteria.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Step 3: The normalized direct relation/influence matrix $\boldsymbol{N}$ can be obtained Eq. (1), in which all principal diagonal elements are equal to zero.

$$\boldsymbol{N} = z\boldsymbol{A}. \tag{1}$$

where $z = \min\left\{\max_i \sum_{j=1}^{n} a_{ij}, \max_j \sum_{i=1}^{n} a_{ij}\right\}$, In this case, $\boldsymbol{N}$ is called the normalized matrix,

Step 4: Then, the total influence-related matrix $\boldsymbol{T}$ can be obtained using Eq. (2), where $\boldsymbol{I}$ stands for the identity matrix.

$$\boldsymbol{T} = \boldsymbol{N} + \boldsymbol{N}^2 + \ldots + \boldsymbol{N}^h = \boldsymbol{N}(\boldsymbol{I} - \boldsymbol{N})^{-1}, \text{ when } \lim_{h\to\infty} \boldsymbol{N}^h = [0]_{n\times n} \tag{2}$$

where $\boldsymbol{T}$ is a total influence-related matrix; $\boldsymbol{N}$ is a direct influence matrix and $\boldsymbol{N} = [x_{ij}]_{n\times n}$; $\lim_{h\to\infty}\left(\boldsymbol{N}^2 + \cdots + \boldsymbol{N}^h\right)$ stands for indirect influence matrix and $0 \le \sum_{j=1}^{n} x_{ij} \le 1$ or $0 \le \sum_{i=1}^{n} x_{ij} \le 1$, and at least one column sum $\sum_j x_{ij}$ or one row sum $\sum_i x_{ij}$, but not all, equals one; then $\lim_{h\to\infty} N^h = [0]_{n\times n}$ can be guaranteed.

The (i, j) element t_{ij} of total influence-related matrix $\boldsymbol{T} = [t_{ij}]_{n\times n}$ denotes the direct and indirect influences of criterion *i* on criterion *j*.

Step 5: The row and column sums are separately denoted as vector ***r*** and vector ***c*** within the total influence-related matrix $\boldsymbol{T}$ as follows:

$$\boldsymbol{T} = [t_{ij}], \quad i, j \in \{1, 2, ..., n\} \tag{3}$$

$$\boldsymbol{r} = [r_i]_{n\times 1} = \left[\sum\nolimits_{j=1}^{n} t_{ij}\right]_{n\times 1} \tag{4}$$

$$\boldsymbol{c} = [c_j]_{n\times 1} = \left[\sum\nolimits_{i=1}^{n} t_{ij}\right]'_{1\times n} \tag{5}$$

where the ***r*** and ***c*** vectors denote the sums of the rows and columns, respectively.

Step 6: Suppose r_i denotes the row sum of the i^{th} row of total influence-related matrix ***T***. Then, r_i is the sum of the influences dispatching from criterion *i* to all other criteria, both directly and indirectly. Suppose that c_j denotes the column sum of the j^{th} column of total influence-related matrix $\boldsymbol{T}$. Then, c_j is the sum of the influences that criterion *j* is received influence from all other criteria. Furthermore, when $i = j$, $(r_i + c_i)$ is the degree of total influence in criterion i (i.e., criterion i is influences all other criteria and is received influence from all other criteria); $(r_i - c_i)$ is the degree of different influence. If $(r_i - c_i)$ is positive, then criterion i primarily is dispatching influence upon the strength of all other criteria; and if $(r_i - c_i)$ is negative, then factor i primarily is received influence from all other criteria (Huang, Shyu, & Tzeng, 2007; Liou, Tzeng, & Chang, 2007; Liou & Tzeng, 2011; Chen & Tzeng, 2011; Tzeng & Huang, 2012; Liu, Tzeng & Lee, 2012).

3.2 DANP for Finding the Influential Weights in Each Criterion

Service quality of cloud computing includes six criteria: (1) privacy and security (C_1), continuity of service (C_2), service migration (C_3), scalable storage (C_4), scaling quickly (C_5), organizational change (C_6). These criteria are not independent, are inter-dependent and feedback in influential relationship problems of each other, and how we know the influential weights? So we cannot use traditional AHP and ANP methods for solving these problems.

New ANP can be easily to be solved these problems by using DEMATEL technique. We can use the basic concept of ANP (Saaty, 1996) by using DEMATEL technique to build the normalized super-matrix for finding the relative influence-weights; in these processes we called DANP (DEMATEL-based ANP) as following steps in general form:

Step 1: Find the normalized matrix $\boldsymbol{T}_c^{nor}$ by dimensions/clusters from total influence-related matrix $\boldsymbol{T}_c$ by criteria. As shown in Eqs.(6-7).

$$\boldsymbol{T}_c = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} & D_1 & & D_j & & D_m \\ & c_{11}\cdots c_{1n_1} & \cdots & c_{j1}\cdots c_{jn_j} & \cdots & c_{m1}\cdots c_{mn_m} \end{array} \\ \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1n_1} \\ c_{i1} \\ c_{i2} \\ \vdots \\ c_{in_i} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mn_m} \end{array} \begin{bmatrix} \boldsymbol{T}_c^{11} & \cdots & \boldsymbol{T}_c^{1j} & \cdots & \boldsymbol{T}_c^{1m} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{i1} & \cdots & \boldsymbol{T}_c^{ij} & \cdots & \boldsymbol{T}_c^{im} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{m1} & \cdots & \boldsymbol{T}_c^{mj} & \cdots & \boldsymbol{T}_c^{mm} \end{bmatrix} \end{array} \tag{6}$$

$$\boldsymbol{T}_c^{nor} = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} & D_1 & & D_j & & D_m \\ & c_{11}\cdots c_{1n_1} & \cdots & c_{j1}\cdots c_{jn_j} & \cdots & c_{m1}\cdots c_{mn_m} \end{array} \\ \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1n_1} \\ c_{i1} \\ c_{i2} \\ \vdots \\ c_{in_i} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mn_m} \end{array} \begin{bmatrix} \boldsymbol{T}_c^{nor_{11}} & \cdots & \boldsymbol{T}_c^{nor_{1j}} & \cdots & \boldsymbol{T}_c^{nor_{1m}} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{nor_{i1}} & \cdots & \boldsymbol{T}_c^{nor_{ij}} & \cdots & \boldsymbol{T}_c^{nor_{im}} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{T}_c^{nor_{m1}} & \cdots & \boldsymbol{T}_c^{nor_{mj}} & \cdots & \boldsymbol{T}_c^{nor_{mm}} \end{bmatrix} \end{array} \tag{7}$$

Step 2: Build an unweighted supermatrix $\boldsymbol{W}_c$. Then, the total influence-related matrix is normalized into a supermatrix according to the interdependence on relationships of dimensions/clusters to obtain an unweighted supermatrix $\boldsymbol{W}_c$ as shown in Eq.(8).

$$\boldsymbol{W}_c = \left(\boldsymbol{T}_c^{nor}\right)' = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} & D_1 & & D_j & & D_m \\ & c_{11}\cdots c_{1n_1} & \cdots & c_{j1}\cdots c_{jn_j} & \cdots & c_{m1}\cdots c_{mn_m} \end{array} \\ \begin{array}{c} c_{11} \\ \vdots \\ c_{1n_1} \\ c_{i1} \\ \vdots \\ c_{in_i} \\ c_{m1} \\ \vdots \\ c_{mn_m} \end{array} \begin{bmatrix} \boldsymbol{W}_c^{11} & \cdots & \boldsymbol{W}_c^{i1} & \cdots & \boldsymbol{W}_c^{m1} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}_c^{1j} & \cdots & \boldsymbol{W}_c^{ij} & \cdots & \boldsymbol{W}_c^{mj} \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{W}_c^{1m} & \cdots & \boldsymbol{W}_c^{im} & \cdots & \boldsymbol{W}_c^{mm} \end{bmatrix} \end{array} \tag{8}$$

Unweighted supermatrix $\boldsymbol{W}_c$ is the matrix transposed from $\boldsymbol{T}_c^{nor}$ (basic concept from the ANP by Saaty (1996), but different from the traditional ANP). If a blank or 0 is shown in the matrix, this means that the dimensions/criteria are independent, and in the same way, a submatrix of $\boldsymbol{W}_c$ can be obtained (Eq.8).

Step 3: Find the normalized total-influential matrix $\boldsymbol{T}_D^{nor}$. The total influence-related matrix $\boldsymbol{T}_D$ needs to be normalised by dividing it by the following formula:

$$t_D^i = \sum_{j=1}^{m} t_D^{ij}$$

$$\boldsymbol{T}_D = \begin{bmatrix} t_D^{11} & \cdots & t_D^{1j} & \cdots & t_D^{1m} \\ \vdots & & \vdots & & \vdots \\ t_D^{i1} & \cdots & t_D^{ij} & \cdots & t_D^{im} \\ \vdots & & \vdots & & \vdots \\ t_D^{m1} & \cdots & t_D^{mj} & \cdots & t_D^{mm} \end{bmatrix} \begin{matrix} \\ \\ \rightarrow \sum_{j=1}^{m} t_D^{1j} = t_D^1 \\ \\ \rightarrow \sum_{j=1}^{m} t_D^{ij} = t_D^i \\ \\ \rightarrow \sum_{j=1}^{m} t_D^{mj} = t_D^m \end{matrix} \tag{9}$$

Thus, the total influence-related matrix can be normalized and shown as $\boldsymbol{T}_D^{nor}$. Then the sum of each row can be defined as being $t_D^i = \sum_{j=1}^{m} t_D^{ij}$, where $i = 1, ..., m$, and $\boldsymbol{T}_D$ can be normalized by the rows of sums by dividing the elements of dimension in each row by the sum of the row to obtain the following. Therefore, a total influence-related matrix can be normalized and shown as $\boldsymbol{T}_D$. $\boldsymbol{T}_D^{nor} = \left[t_D^{ij} / t_D^i \right]_{m \times m}$, like Eq.(10). Afterwards, each row of the normalized $\boldsymbol{T}_D^{nor}$ can be summarized as equal one, so that $\sum_{j=1}^{m} t_D^{nor_{ij}} = 1$.

$$\boldsymbol{T}_D^{nor} = \begin{bmatrix} t_D^{11} / t_D^1 & \cdots & t_D^{1j} / t_D^1 & \cdots & t_D^{1m} / t_D^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_D^{i1} / t_D^i & \cdots & t_D^{ij} / t_D^i & \cdots & t_D^{im} / t_D^i \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_D^{m1} / t_D^m & \cdots & t_D^{mj} / t_D^m & \cdots & t_D^{mm} / t_D^m \end{bmatrix}$$

$$= \begin{bmatrix} t_D^{nor_{11}} & \cdots & t_D^{nor_{1j}} & \cdots & t_D^{nor_{1m}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_D^{nor_{i1}} & \cdots & t_D^{nor_{ij}} & \cdots & t_D^{nor_{im}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_D^{nor_{m1}} & \cdots & t_D^{nor_{mj}} & \cdots & t_D^{nor_{mm}} \end{bmatrix} \tag{10}$$

where $\boldsymbol{T}_D^{nor} = \left[t_D^{nor_{ij}} \right] = \left[t_D^{ij} / t_D^i \right]$, i.e., $\boldsymbol{T}_D^{nor_{ij}} = \left[t_D^{ij} / t_D^i \right]$. Thus, this study adopts a normalized total-influential matrix $\boldsymbol{T}_D^{nor}$ (hereafter abbreviated to "the normalized matrix").

Step 4: Find the influential weights of the DANP. The total influence-related matrix $\boldsymbol{T}_c$ needs to be normalized by dividing the dimension/cluster (Eq. (6)), so $\boldsymbol{T}_c$ is normalized by summarizing the row by dimensions/clusters to get $\boldsymbol{T}_c^{nor}$. An

unweighted super-matrix $\boldsymbol{W}_c$ can be obtained by transposing $\boldsymbol{T}_c^{nor}$, i.e. $\boldsymbol{W}_c = \left(\boldsymbol{T}_c^{nor}\right)'$. Using (Eq. (11)), a weighted super-matrix $\boldsymbol{W}_c^{\ *}$ (improving the traditional ANP by using equal weights to make it appropriate for the real world) can be obtained by the production of $\boldsymbol{T}_D^{nor}$ and $\boldsymbol{W}_c$, i.e. $\boldsymbol{W}_c^{\ *} = \boldsymbol{T}_D^{nor} \times \boldsymbol{W}_c$ (Eq. (11)). This demonstrates that these influential level values are the basis of normalization to determine a weighted super-matrix.

$$\boldsymbol{W}_c^{\ *} = \boldsymbol{T}_D^{nor} \times \boldsymbol{W}_c = \begin{bmatrix} t_D^{nor_{11}} \times \boldsymbol{W}_c^{\ 11} & \cdots & t_D^{nor_{i1}} \times \boldsymbol{W}_c^{\ i1} & \cdots & t_D^{nor_{m1}} \times \boldsymbol{W}_c^{\ m1} \\ \vdots & & \cdots & & \vdots \\ t_D^{nor_{1j}} \times \boldsymbol{W}_c^{\ 1j} & \cdots & t_D^{nor_{ij}} \times \boldsymbol{W}_c^{\ ij} & \cdots & t_D^{nor_{mj}} \times \boldsymbol{W}_c^{\ mj} \\ \vdots & & \vdots & & \vdots \\ t_D^{nor_{1m}} \times \boldsymbol{W}_c^{\ 1m} & \cdots & t_D^{nor_{im}} \times \boldsymbol{W}_c^{\ im} & \cdots & t_D^{nor_{mm}} \times \boldsymbol{W}_c^{\ mm} \end{bmatrix} \quad (11)$$

Step 5: Obtain the DANP. Limit the weighted super-matrix by raising it to a sufficiently large power q until it has converged and becomes a long-term stable super-matrix to obtain global priority vectors, or the so-called DANP, in influential weights by $\lim_{q \to \infty} (\boldsymbol{W}_c^{\ *})^q$. Therefore, the DANP in influential weights $\boldsymbol{w} = (w_1, \ldots, w_j, \ldots, w_n)$ can be obtained by Eq.(11).

3.3 Grey Relational Analysis for Evaluation and Improvement

Since Deng (1986) proposed Grey theory, related models have been developed and applied to MCDM problems. Similar to fuzzy set theory, Grey theory is a feasible mathematical means that can be used to deal with systems analysis characterized by inadequate information. Fields covered by the Grey theory include systems analysis, data processing, modeling, prediction, decision-making, and control engineering (Deng, 1985; Deng, 1988; Deng, 1989; Tzeng, & Tasur, 1994). In this section, we briefly review some relevant definitions and the calculation process for the Grey Relation Model. This research modified the definitions by Chiou and Tzeng (2001) and produced the definitions indicated below. GRA is used to determine the relationship between two sequences of stochastic data in a Grey system. The procedure bears some similarity to pattern recognition technology. One sequence of data is called the 'reference pattern' or 'reference sequence,' (in our paper we set the reference sequence as aspiration level, i.e., if the measures of scaling are from zero to ten (0,1,2,…,10), we set zero (0) is the worst value and ten (10) is aspiration level; then how can improve cloud computing service quality for closing the aspiration level; it is key issue in big contribution of this research) and the correlation between the other sequence and the reference sequence is to be identified (Tzeng, & Tasur, 1994; Deng, 1986; Mon, Tzeng, Lu, 1995; Wu, Deng, & Wen, 1996). The processes are shown as follows:

Step 1: Find the performance matrix in alternatives for improvement strategies. The performance matrix Eq. (12) of each criterion in each alternative for improvement strategies can be obtained by questionnaires using scaling measures from 0 points (for complete dissatisfaction) to 10 points (for the best satisfaction), so the aspiration level can be set at ten and the worst value at zero. Therefore, in this research, ten is set as the aspiration level and zero as the worst value, which differs from traditional approach settings.

Step 2: Build the initial relationship matrix.
Let the initial relationship matrix be a $m \times n$ matrix, where there are m alternatives for improvement strategies and n criteria, obtained by surveying the relationships as follows.

$$\begin{array}{lccccc} & & & \boxed{\text{Criteria}} & & \\ \text{Alternatives} & c_1 & \cdots & c_j & \cdots & c_n \\ & w_1 & \cdots & w_j & \cdots & w_n \\ x_1 & \left[\begin{array}{ccccc} x_1(1) & \cdots & x_1(j) & \cdots & x_1(n) \end{array}\right. & & & & \\ \vdots & \vdots & & \vdots & & \vdots \\ x_k & x_k(1) & \cdots & x_k(j) & \cdots & x_k(n) \\ \vdots & \vdots & & \vdots & & \vdots \\ x_m & \left.\begin{array}{ccccc} x_m(1) & \cdots & x_m(j) & \cdots & x_m(n) \end{array}\right] & & & & \\ \text{Aspiration level } x^* & x^*(1) & \cdots & x^*(j) & \cdots & x^*(n) \end{array} \tag{12}$$

Step 3: Find the coefficients of each grey relation.

Therefore, the coefficients of each grey relation for aspiration level are

$$:\gamma(x^*(j), x_k(j)) = \frac{\min\limits_k \min\limits_j |x^*(j) - x_k(j)| + \varsigma \max\limits_k \max\limits_j |x^*(j) - x_k(j)|}{|x^*(j) - x_k(j)| + \varsigma \max\limits_k \max\limits_j |x^*(j) - x_k(j)|} \tag{13}$$

Step 4: Find the grade of grey relation.
Then, the grade (degree) of grey relation is obtained, the larger the better (shown as the grade of coming close to the aspired values):

$$\gamma(x^*, x_k) = \sum_{j=1}^{n} w_j \gamma(x^*(j), x_k(j)), \tag{14}$$

where j^{th} denotes j^{th} criterion and $j = 1, 2, ..., n;$ k^{th} denotes k^{th} alternative and $k = 1, 2, ..., m;$ the influence weight w_j can be obtained by DANP.

4 Empirical Case Study

Based on the literature review results, the six barriers for introducing cloud computing into a mobile telecommunications service provider are concluded by the authors.

4.1 Criteria Derivation

First, privacy and security implies that customers have concerned on their privacy and data security than traditional hosting service. Second, the continuity of service refers to the factors that may negatively affect the continuity of cloud computing such as Internet problems, power cut-off, service disruption and system bugs. Third, the service migration means that no regularity organization have reached the agreement on the standardization of the cloud computing's external interface (Qian, Luo, Du, Guo, 2009). Fourth, the scalable storage implies to create a storage system would not only meet these needs but combine them with the cloud advantages of scaling arbitrarily up and down on-demand, as well as meeting programmer expectations in regard to resource management for scalability, data durability, and high availability. The fifth is scaling quickly. The opportunity is then to automatically scale quickly up and down in response to load in order to save money, but without violating service level agreements (Vaquero, Merino, Caceres, & Lindner, 2009). The last one is the organizational change. After introducing cloud computing into the enterprises, the traditional management does not apply to those enterprises. Therefore, the change of organization is necessary.

In order to solve the problems that mentioned before and to promote the service quality in telecommunication, several possible solutions are introduced. For reliability, adequate funding and new education or training should be guaranteed. For tangibles, to become a platform may be an idea. For responsiveness, the enterprises have to establish customer relationship well. For assurance, some basic infrastructures have to be well-prepared, so better practices for handling dynamic workloads and dynamic capacity management become very important. For empathy, an enterprise need to understand how the customers feel, and then trying to improve any problems.

4.2 The Result by DEMATEL, DANP and GRA

By using the DEMATEL, the relationship between the six criteria can be constructed and shown in the Figure 2. We also find the threshold is 0.9523. The item whose value is over the threshold is influential in this research. Furthermore, the authors draw a relation diagram to show the relationship between every criterion.

Privacy and security (C1) is influenced by nothing. Continuity of service (C2) is influenced by scaling quickly. Service migration (C3) is influenced by nothing. Scalable storage (C4) is influenced by service migration, scaling quickly, and organizational change. Scaling quickly (C5) is influenced by continuity of service, service migration, scalable storage, and organizational change. Organizational change (C6) is influenced by privacy and security and service migration.

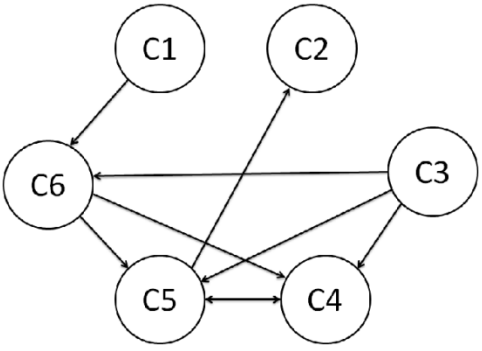


Fig. 2 The relationship between the six criteria

Then we calculate the weights by ANP, and find that scalable storage and scaling quickly have the highest weight. The weights are 0.499 and 0.342. The outcome means the two criteria are the important reason for improving cloud computing in the business.

Dynamic storage in cloud computing is a main part in this issue that can be explained. Organizational change is also important to an enterprise which introduced cloud computing in the organization. It got 0.293 in ANP. It is easy to understand when a company uses a new way to run its business, the establishment of the department and the way of management should be changed. The weights are shown in Table 1.

Table 1 Weights versus Each Barrier

Criteria	Weights
privacy and security	0.265
continuity of service	0.126
service migration	0.138
scalable storage	0.499
scaling quickly	0.342
organizational change	0.293

In Table2, the result of GRA is showed, and we can see that new education and training have the highest scores. It got 0.860. The second and the third are expertise in dynamic capacity management and vastly superior economics. They got 0.758 and 0.696. According to the result, some discussions can be had in the following chapter.

Table 2 The result of Grey Relational Grades

Strategies	Grey
Vastly superior economics	0.696
Leverage existing investment	0.680
Defend a franchise	0.447
Leverage customer relationships	0.687
Become a platform	0.677
Better practices for handling dynamic workloads	0.544
New education and training	0.860
Dynamic capacity management	0.758

5 Discussion

By analyzing these data, dynamic storage and organizational change are the most serious problems when we are talking about introducing cloud computing into the telecommunication mobile providers that can be found. If the users become more and more, the providers must need a dynamic space to storage the customs' communication data. The providers also need the flexible storage to save money and make the function smoothly. In this way, the GRA outcome can be combined. It showed us that expertise in dynamic capacity management is one of very essential elements in using cloud computing in improving the function in telecom mobile providers.

Furthermore, when organization changes, the management approach usually happens concurrently. Now, a new management approach has been introduced. And thus, the organization must be changed. That makes education and training essential, like the GRA outcome demonstrated. Any employee should have the ability to face the troubles and deals with the new system in the organization. Therefore, good education and training courses should be established to let the employees realize the new technology so that they can reach a good performance.

The GRA results demonstrate that vastly superior economics is important exactly. Any improvement, education, new hardware, and change all need money. In general, any enterprise which can control much money, it has the more opportunities to be successful.

In this paper, the primary barriers being derived were where the telecom mobile providers should introduce cloud computing and the solutions.

6 Conclusion

Cloud computing are changing the Internet world. Many enterprises also have introduced cloud computing to improve their performance. In this research, organizational changing, and scalable storage_which are found by the authors by using a new hybrid MCDM model are the two problems that a mobile communications service provider must solve; expertise in dynamic capacity management and education and training are the two suitable ways to deal with these problems. As same

as the interview to a manager of one mobile telecommunication company, this is exactly a serious problem that they have to improve to decrease the coat and strengthen the service quality.

References

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G.: Above the Clouds: A Berkeley View of Cloud Computing. Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, C.A. (2009)

[2] Chen, C.H., Tzeng, G.H.: Creating the aspired intelligent assessment systems for teaching materials. Expert Systems with Applications 38(10), 12168–12179 (2011)

[3] Chen, F.H., Hsu, T.S., Tzeng, G.H.: A balanced scorecard approach to establish a performance evaluation and relationship model for hot spring hotels based on a hybrid MCDM model combining DEMATEL and ANP. International Journal of Hospitality Management 30(4), 908–932 (2011)

[4] Chiou, H.K., Tzeng, G.H.: Fuzzy hierarchical evaluation with Grey relation model of green engineering for industry. International Journal of Fuzzy System 3(3), 466–475 (2001)

[5] Chiu, Y.J., Chen, H.C., Tzeng, G.H., Shyu, J.Z.: Marketing strategy based on customer behavior for the LCD-TV. International Journal of Management and Decision Making 7(2/3), 143–165 (2006)

[6] Deng, J.L.: Fundamental Methods of Grey Systems. Huazhoug University of Science and Technology, Wuhan (1985)

[7] Deng, J.L.: Grey Forecasting and Decision. Huazhong University of Science and Technology Press, Wuhan (1986)

[8] Deng, J.L.: Grey System Book. Science and Technology Information Services, Windsor (1988)

[9] Deng, J.L.: Introduction of grey theory. The Journal of Grey System 1(1), 1–24 (1989)

[10] Gabus, A., Fontela, E.: World problems an invitation to further thought within the framework of DEMATEL. Battelle Geneva Research Centre, Geneva (1972)

[11] Headley, D.E., Miller, S.J.: Measuring service quality and its relationship to future consumer behavior. Journal of Health Care Marketing 13(4), 32 (1993)

[12] Ho, W.R., Tsai, C.L., Tzeng, G.H., Fang, S.K.: Combined DEMATEL technique with a novel MCDM model for exploring portfolio selection based on CAPM. Expert Systems with Applications 38(1), 16–25 (2011)

[13] Huang, C.Y., Shyu, J.Z., Tzeng, G.H.: Reconfiguring the innovation policy portfolios for Taiwan's SIP Mall industry. Technovation 27(12), 744–765 (2007)

[14] Huang, J.J., Tzeng, G.H., Ong, C.S.: Multidimensional data in multidimensional scaling using the analytic network process. Pattern Recognition Letters 26(6), 755–767 (2005)

[15] Huang, C.Y., Tzeng, G.H., Ho, W.R.: System on chip design service e-business value maximization through a novel MCDM framework. Expert Systems with Applications 38(7), 7947–7962 (2011)

[16] Jeng, D.J.F., Tzeng, G.H.: Social influence on the use of clinical decision support systems: revisiting the unified theory of acceptance and use of technology by the fuzzy DEMATEL technique. In: Computers & Industrial Engineering (in Press, 2012)
[17] K-Hosseini, A., Sommerville, I., Sriram, I.: Research Challenges for Enterprise Cloud Computing, Cornell University (2010)
[18] Kuan, M.J., Tzeng, G.H., Hsiang, C.C.: Exploring the quality assessment system for new product development process by combining DANP with MCDM model. International Journal of Innovative Computing, Information and Control (Special Issue) 8(8) (in press, 2012)
[19] Lee, Y.T., Wu, W.W., Tzeng, G.H.: Combining the DEMATEL with the ANP and ZOGP for selecting IT projects. Information Science (2007) (forthcoming)
[20] Lin, C.J., Wu, W.W.: A causal analytical method for group decision-making under fuzzy environment. Expert Systems with Applications 34(1), 205–213 (2008)
[21] Liou, J.J.H., Tzeng, G.H., Chang, H.C.: Airline safety measurement using a hybrid model. Air Transport Management 13(4), 243–249 (2007)
[22] Liou, J.J.H., Tzeng, G.H.: A hybrid ANP model in fuzzy environments for strategic alliance partner selection in the airline industry. Applied Soft Computing 11(4), 3515–3524 (2011)
[23] Liou, J.J.H., Tzeng, G.H., Hsu, C.C., Yeh, W.C.: Reply to "comment on using a modified grey relation method for improving airline service quality. Tourism Management (in press, 2012)
[24] Liu, C.H., Gwo-Hshiung Tzeng, G.H., Lee, M.H.: Improving tourism policy implementation - the use of hybrid MCDM models. Tourism Management 33(2), 239–488 (2012)
[25] Liu, C.H., Tzeng, G.H., Lee, M.H.: Strategies for improving cruise product sales in the travel agency- using hybrid MCDM models. The Service Industry Journal (in press, 2012)
[26] Liou, J.J.H., Tsai, C.Y., Lin, R.H., Tzeng, G.H.: A modified VIKOR multiple-criteria decision method for improving domestic airlines service quality. Journal of Air Transport Management 17(2), 57–61 (2011)
[27] Mell, P., Grance, T.: NIST Definition of Cloud Computing. National Institute of Standards and Technology (2009)
[28] Mon, D.L., Tzeng, G.H., Lu, H.C.: Grey decision making in weapon system evaluation. Journal of Chung Chen Institute of Technology 26(1), 73–84 (1995)
[29] Myoken, Y.: Cloud Computing in Japan, ICT & Engineering Team (2009)
[30] Parasuraman, A., Berry, L.L., Zeithaml, V.A.: Understanding Customer Expectations of Service. Sloan Management Review 32 (1991)
[31] Ou Yang, Y.P., Shieh, H.M., Tzeng, G.H.: A IKOR Technique Based on DEMATEL and ANP for Information Security Risk Control Assessment. Information Sciences (in press, 2012)
[32] Parkhill, D.F.: The challenge of the computer utility, p. 246. Addison-Wesley Pub. Co. (1966)
[33] Qian, L., Luo, Z., Du, Y., Guo, L.: Cloud Computing: An Overview. In: Jaatun, M.G., Zhao, G., Rong, C. (eds.) Cloud Computing. LNCS, vol. 5931, pp. 626–631. Springer, Heidelberg (2009)
[34] Saaty, R.W.: The Analytic Hierarchy Process (AHP) for Decision Making and The Analytic Network Process (ANP) for Decision Making with Dependence and Feedback. Creative Decisions Foundation, Pittsburgh (2003)

[35] Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process. RWS Publication, Pittsburgh (1996)
[36] Saaty, T.L.: Decision Making With Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburgh (1996)
[37] Saaty, T.L.: Fundamentals of the Analytic Network Process. In: Proceedings of International Symposium on Analytical Hierarchy Process, Japan, Kobe (1999)
[38] Saaty, T.L.: Fundamentals of the analytic network process - dependence and feedback in decision-making with a single network. Journal of Systems Science and Systems Engineering 13(2), 71–91 (2004)
[39] Saaty, T.L.: Theory and Applications of the Analytic Network Process. RWS Publications, Pittsburg (2005)
[40] Shen, Y.C., Lin, G.T.R., Tzeng, G.H.: Combined DEMATEL techniques with novel MCDM for the organic light emitting diode technology selection. Expert Systems with Applications 38(3), 1468–1481 (2011)
[41] Staten, J.: Is Cloud Computing Ready For The Enterprise? Forrester Research, Inc. (2008)
[42] Tamura, M., Nagata, H., Akazawa, K.: Extraction and systems analysis of factors that prevent safety and security by structural models. In: Proceedings of the 41st SICE Annual Conference, Japan, Osaka, vol. 3, pp. 1752–1759 (2002)
[43] Tzeng, G.H., Tasur, S.H.: The Multiple Criteria Evaluation of Grey Relation Model. The Journal of Grey System 6(2), 87–108 (1994)
[44] Tzeng, G.H., Chiang, C.H., Li, C.W.: Evaluating intertwined effects in e-learning programs: a novel hybrid MCDM model based on factor analysis and DEMATEL. Expert Systems with Applications 32(4), 1028–1044 (2007)
[45] Tzeng, G.H., Huang, C.Y.: Combined DEMATEL technique with hybrid MCDM methods for creating the aspired intelligent global manufacturing & logistics systems. Annals of Operations Research (2011) (in press)
[46] Vaquero, L.M., Merino, L.R., Caceres, J., Lindner, M.: A break in the clouds: Towards a cloud definition. ACM SIGCOMM Computer Communication Review 39(1), 50 (2009)
[47] Vouk, M.A.: Cloud computing – issues, research and implementations. Journal of Computing and Information Technology 16(4), 235–246 (2008)
[48] Wang, Y.L., Tzeng, G.H.: Brand marketing for creating brand value based on a MCDM model combining DEMATEL with ANP and VIKOR methods. Expert Systems with Applications 39(5), 5600–5615 (2012)
[49] Wu, H.S., Deng, J.L., Wen, K.L.: Introduction of Grey A (1996)
[50] Wu, W.W., Lee, Y.T.: Developing global managers' competencies using the fuzzy DEMATEL method. Expert Systems with Applications 32(2), 499–507 (2007)
[51] Yang, J.L., Tzeng, G.H.: An integrated MCDM technique combined with DEMATEL for a novel cluster-weighted with ANP method. Expert Systems with Applications 38(3), 1417–1424 (2011)
[52] Yanosky, R.: From Users to Choosers: The Cloud and the Changing Shape of Enterprise Authority. In: Katz, R.N. (ed.) The Tower and the Cloud, pp. 126–136. EDUCAUSE (2008)

Author Index

Subject Index