Advances in Intelligent Systems and Computing 443

Bing-Yuan Cao
Pei-Zhuang Wang
Zeng-Liang Liu
Yu-Bin Zhong *Editors*

International Conference on Oriental Thinking and Fuzzy Logic

Celebration of the 50th Anniversary in the era of Complex Systems and Big Data

Springer

Advances in Intelligent Systems and Computing

Volume 443

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/11156

Bing-Yuan Cao · Pei-Zhuang Wang
Zeng-Liang Liu · Yu-Bin Zhong
Editors

International Conference on Oriental Thinking and Fuzzy Logic

Celebration of the 50th Anniversary in the era of Complex Systems and Big Data

Springer

Editors
Bing-Yuan Cao
Guangzhou Vocational College of Science and Technology
Guangzhou
China

and

Guangzhou University
Guangzhou
China

Pei-Zhuang Wang
West Texas A&M University
Canyon, TX
USA

and

Liaoning Technical University
Fuxin
China

Zeng-Liang Liu
Foreign Affairs Office
National Defense University
Beijing
China

Yu-Bin Zhong
School of Mathematics and Information Science
Guangzhou University
Guangzhou
China

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-30873-9 ISBN 978-3-319-30874-6 (eBook)
DOI 10.1007/978-3-319-30874-6

Library of Congress Control Number: 2016934435

Printed on acid-free paper

This Springer imprint is published by SpringerNature
The registered company is Springer International Publishing AG Switzerland

Preface

Oriental Thinking and Fuzzy Logic in Dalian, China, an international conference, was held during August 17–20, 2015, to celebrate the 50th anniversary of Fuzzy Sets. The honorary chair for this conference was Prof. L.A. Zadeh, the founder of fuzzy sets theory, who has guided an information revolution, and constructed a great bridge between qualitative and quantitative.

The conference focused on six main topics as follows: fuzzy information processing; fuzzy engineering; Internet and big data applications; factor space and factorial neural networks; information granulation and granular computing; extension and innovation methods. Here, topic three, the theory of factor space was initiated by Prof. Pei-Zhuang Wang with the oriental thinking. And extension topic six, is a new field of the disciplinary initiated by Prof. Wen Cai, who achieved innovation facing a problem where impossible cases seem to be possible.

There were 15 plenary talks in the conference including Wen Cai, fuzzy logic and extenics; Y.X. Chen, inter-definability and application of fuzzy logic operators; I. Dzitac, fuzzy logic and artificial intelligence; J.L. Feng, theory of meta-synthetic wisdom based on fusion of qualitative, quantitative and imagery operations; J.F. Gu, system science and Chinese medicine; Ouyang He, a mathematical foundation for factor spaces; Qing He, uncertainty learning; C.F. Huang, an approach checking whether an intelligent internet can be improved into intelligence; D.Y. Li, cognitive physics; Z.L. Liu, factorial neural networks; W. Pedrycz, new frontiers of computing and reasoning with qualitative information: a perspective of granular computing; Germano Resconi, from inconsistent topology to consistent in big data; Yong Shi and Y.J. Tian, uncertainty and big databases; P.Z. Wang, fuzzy sets and factor space; Z.S. Xu, complex information decision making. As a special guest, Mr. H.R. Lin, with his 18-year teaching practice in Shanghai Middle School, introduced his book "Preliminary of Fuzzy Mathematics" for pupils in his schools.

Apart from the organized speeches, we much appreciated the articles from individuals with natural interest and deep friendship toward Prof. L. Zadeh. They developed fuzzy theory along probability representation, rough sets, intuitionistic fuzzy sets, nonlinear Particle Swarm Optimization, ranking method to structure elements and they apply fuzzy theory into recommendation, feature extraction, qualitative mapping, etc. Among all papers presented at the conference, we carefully selected over 60 papers to form this book as assorted appetizers to commemorate the 50th anniversary of fuzzy sets from the Dalian conference.

Finally, we thank the publisher, Springer, for publishing the proceedings as Advance in Intelligent and Soft Computing.

December 2015

Bing-Yuan Cao
Pei-Zhuang Wang
Zeng-Liang Liu
Yu-Bin Zhong

Editorial Committee

Contents

Part VII Others

Part I
Fuzzy Information Processing

$(\in, \in \vee q_{(\lambda,\mu)})$-Fuzzy Weak Ideal of Complemented Semirings

Zuhua Liao, Chan Zhu, Xiaotang Luo, Xiaoying Zhu, Wangui Yuan and Juan Tong

Abstract In this paper,the notions of generalized fuzzy weak ideal of complemented semiring, $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal of complemented semiring are introduced. By discussing, the two new concepts are found to be equivalent. Furthermore, some fundamental properties of their intersection, union, level sets, homomorphic image and homomorphic preimage are investigated.

Keywords $(\in, \in \vee q_{(\lambda,\mu)})$-Fuzzy weak ideal · Generalized fuzzy weak ideal · Homomorphic image · Homomorphic preimage

1 Introduction

Rosenfeld in 1971 introduced fuzzy sets in the context of group theory and formulated the concept of a fuzzy subgroup of a group [1]. Since then, many researches have extended the concepts of abstract algebra to a fuzzy framework. In 2006, Liao etc. generalized "quasi-coincident with" relation (q) between a fuzzy point and a fuzzy set of Liu to "generalized quasi-coincident with" relation ($q_{(\lambda,\mu)}$) between a fuzzy point and a fuzzy set, and extended Rosenfeld's $(\in, \in)$-fuzzy algebra, Bhakat and Das's $(\in, \in \vee q)$-fuzzy algebra and $(\overline{\in}, \overline{\in} \vee \overline{q}_{(\lambda,\mu)})$-fuzzy algebra to $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy algebra [2] with more abundant hierarchy [3].

Vandiver [4] in 1939 put forward the concept of semiring. The applications of semirings to areas such as optimization theory, graph theory, theory of discrete event dynamical systems, generalized fuzzy computation, automata theory, formal language theory, coding theory and analysis of computer programs have been extensively studied in the literature [5, 6]. Liu [7] introduced fuzzy ideals in a ring.

Z. Liao (✉) · C. Zhu · X. Luo · X. Zhu · W. Yuan · J. Tong
School of Science, Jiangnan University, Wuxi 214122, China
e-mail: liaozuhua57@163.com

Z. Liao · C. Zhu · X. Luo · X. Zhu · W. Yuan · J. Tong
Institute of Intelligent Systems and Network Computing,
Engineering of Internet of Things, Jiangnan University, Wuxi 214122, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_1

Following this definition, Mukherjee and Sen [8, 9]obtained many interesting results in the theory of rings.

On the study of fuzzy semiring, there has been a large number of researches at home and abroad. Feng and Zhan [10] proposed complemented semiring. They put the Boolean algebra as its proper class, and studied the algebraic structure of it.

This paper is the continuation of the above work.

Section 2 of this paper list some necessary preliminaries that support our results. Section 3 is the kernel of the whole paper, which display main results obtained by the authors, including the relationships among generalized fuzzy weak ideal, $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal and level subsets of a fuzzy set and relative properties about the intersection, union, homomorphic image and homomorphic preimage of such generalized fuzzy weak ideal. In section, we make a conclusion and prospect the further study of generalized fuzzy weak ideal.

2 Preliminaries

In this section we recall some basic notions and results which will be needed in the sequel.

Throughout the paper we always consider S as a semigroup.

Definition 1 [11] A semiring S is a structure consisting of a nonempty set S together with two binary operations on S called addition and multiplication (denoted in the usual manner) such that

(1) S together with addition is a semigroup. o is additive identity element;
(2) S together with multiplication is a semigroup.1 is multiplicative identity element;
(3) $a(b + c) = ab + ac, (a + b)c = ac + bc, \forall a, b, c \in S$;
(4) $o \cdot a = a \cdot o = o$.

Definition 2 [10] Assume a is an element of semiring S, if there exists a complement $\overline{a}$ which makes $a\overline{a} = o, a + \overline{a} = 1$, a is called complemented. S is said to be a complemented semiring if every element of S has a complement.

Definition 3 [12] Let S be a semiring. A nonempty subset A of S is said to be a complemented subsemiring of S if A is closed under three binary operations on S:

(1) If $a, b \in A$, then $a + b \in A$;
(2) If $a, b \in A$, then $ab \in A$;
(3) If $a \in A$, then $\overline{a} \in A$.

From now on, we write S and H for complemented semirings.

Definition 4 [13] Let $\alpha, \lambda, \mu \in [0, 1]$ and $\lambda < \mu$, if $A(x) \geq \alpha$, then a fuzzy point x_α is said to belongs to a fuzzy subset A written $x_\alpha \in A$; if $\lambda < \alpha$ and $A(x) + \alpha > 2\mu$, then a fuzzy point x_α is called to be generalized quasi-coincident with a fuzzy subset A, denoted by $x_\alpha q_{(\lambda,\mu)} A$. If $x_\alpha \in A$ or $x_\alpha q_{(\lambda,\mu)} A$, then $x_\alpha \in \vee q_{(\lambda,\mu)} A$.

Definition 5 [12] Let $\alpha, \lambda, \mu \in [0,1]$ and $\lambda < \mu$, A fuzzy subset A of S is called an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy complemented subsemiring of S, if $\forall t, r \in (\lambda, 1]$, $a, b \in S$ satisfy:

(1) If $a_t, b_r \in A$, we have $(a+b)_{t\wedge r} \in \vee q_{(\lambda,\mu)} A$;
(2) If $a_t, b_r \in A$, there exist $(ab)_{t\wedge r} \in \vee q_{(\lambda,\mu)} A$;
(3) If $a_t \in A$, $a_t \in \vee q_{(\lambda,\mu)} A$ holds.

Definition 6 [12] Let $\alpha, \lambda, \mu \in [0,1]$ and $\lambda < \mu$, A is a fuzzy set of S. We call A a generalized fuzzy complemented subsemiring of S if $\forall a, b \in S$ satisfy:

(1) $A(a+b) \vee \lambda \geq A(a) \wedge A(b) \wedge \mu$;
(2) $A(ab) \vee \lambda \geq A(a) \wedge A(b) \wedge \mu$;
(3) $A(\overline{a}) \vee \lambda \geq A(a) \wedge \mu$.

Definition 7 [14] Let $S_i (1 \leq i \leq n)$ be complemented semirings and direct product: $\prod_{1\leq i\leq n} S_i = \{(a_1, a_2, \ldots a_n) | a_i \in S_i\}$. Then $\prod_{1\leq i\leq n} S_i$ is a complemented semiring under the operations as following:

$(a_1, a_2, \ldots a_n) + (b_1, b_2, \ldots b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$;
$(a_1, a_2, \ldots a_n)(b_1, b_2, \ldots b_n) = (a_1 b_1, a_2 b_2, \ldots a_n b_n)$;
$\overline{(a_1, a_2, \ldots a_n)} = (\overline{a_1}, \overline{a_2}, \ldots \overline{a_n})$.

Definition 8 [14] Let $A_i (1 \leq i \leq n)$ be fuzzy subsets of S_i, then a fuzzy set $\prod_{1\leq i\leq n} A_i$ defined as $(\prod_{1\leq i\leq n} A_i)(x_1, x_2, \ldots x_n) = \inf_{1\leq i\leq n} A_i(x_i)$ is called fuzzy direct product.

Theorem 1 [12] *Let A be a fuzzy subset of S, then A is a generalized fuzzy complemented subsemiring of S if and only if A is an* $(\in, \in \vee q_{(\lambda,\mu)})$*-fuzzy complemented subsemiring of S.*

Theorem 2 [12] *Let A be a fuzzy subset of S, then A is a generalized fuzzy complemented subsemiring of S if and only if* $\forall \alpha \in (\lambda, \mu]$*, nonempty* A_α *is a subsemiring of S.*

Theorem 3 [12] *Let A and B be generalized fuzzy complemented subsemirings of S, then* $A \cap B$ *is a generalized fuzzy complemented subsemiring of S.*

Theorem 4 [12] *A subset A of S is a complemented subsemiring of S if and only if* χ_A *is a generalized fuzzy complemented subsemiring of S.*

Theorem 5 [12] *Let* $f : S \to H$ *be a full homomorphism. If A is a generalized fuzzy complemented subsemiring of S, then* $f(A)$ *is a generalized fuzzy complemented subsemiring of H.*

Theorem 6 [12] *Let* $f : S \to H$ *be a homomorphism. If B is a generalized fuzzy complemented subsemiring of H, then* $f^{-1}(B)$ *is a generalized fuzzy complemented subsemiring of S.*

3 $(\in, \in \vee q_{(\lambda,\mu)})$-Fuzzy Completely Prime Ideals

The following if no special instructions, we suppose $\lambda, \mu \in [0, 1]$ and $\lambda < \mu$. Firstly, the definitions of weak idea of complemented semiring generalized fuzzy weak ideal and $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal are given.

Definition 9 Let A be a complemented subsemiring of S, then A is said to be a weak idea of complemented semiring S if for all $x \in A, y \in S, xy + yx \in A$.

Definition 10 Let A be a generalized fuzzy complemented subsemiring of S, then A is called a generalized fuzzy weak ideal of S if $A(xy + yx) \vee \lambda \geq A(x) \wedge \mu$. $\forall x, y \in S$.

Definition 11 Let A be an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy complemented subsemiring of S. Then A is called an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal of S, if $\alpha \in (\lambda, 1], y \in S, x_\alpha \in A$ imply $(xy + yx)_\alpha \in \vee q_{(\lambda,\mu)} A$.

By the research on the relationships among generalized fuzzy complemented subsemiring, $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal, and level subsets of a fuzzy set, we obtain the following result.

Theorem 7 *Let A be a fuzzy subset of S, then the following conditions are equivalent:*

(1) A is an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal of S;
(2) A is a generalized fuzzy weak ideal of S;
(3) $\forall \alpha \in (\lambda, \mu]$, nonempty set A_α is a weak ideal of S.

Proof (1) $\Rightarrow$ (2):

From Theorem 1, we know A is an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy complemented subsemiring of S thus A is a generalized fuzzy complemented subsemiring of S.

Next we prove A is a generalized fuzzy weak ideal of A. Assume that there exist $x_0, y_0 \in S$ such that $A(x_0y_0 + y_0x_0) \vee \lambda < A(x_0) \wedge \mu$. choose α such that $A(x_0y_0 + y_0x_0) \vee \lambda < \alpha < A(x_0) \wedge \mu$, then $A(x_0y_0 + y_0x_0) < \alpha, A(x_0) > \alpha$ and $\lambda < \alpha < \mu$, so $(x_0)_\alpha \in A$. Since A is an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal of S, thus $(x_0y_0 + y_0x_0)_\alpha \in \vee q_{(\lambda,\mu)} A$. But $A(x_0y_0 + y_0x_0) + \alpha < \alpha + \alpha < 2\mu$, a contradiction.

So A is a generalized fuzzy weak ideal of S.

(2) $\Rightarrow$ (1)

It is easy to prove that A is an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy complemented subsemiring of S. $\forall x, y \in S, \alpha \in (\lambda, 1]$, if $x_\alpha \in A$, then $A(x) \geq \alpha$. Since A is a generalized fuzzy weak ideal of S, then $A(xy + yx) \vee \lambda \geq A(x) \wedge \mu \geq \alpha \wedge \mu$.

Case 1: If $\alpha > \mu$, then $A(xy + yx) \vee \lambda \geq \mu$. By $\lambda \leq \mu$, so $A(xy + yx) \geq \mu$. Then $A(xy + yx) + \alpha \geq \mu + \alpha > 2\mu$, i.e.$(xy + yx)_\alpha q_{(\lambda,\mu)} A$.

Case 2: If $\alpha \leq \mu$, then we can obtain that $A(xy + yx) \geq \alpha$, i.e.$(xy + yx)_\alpha \in A$. So $(xy + xy)_\alpha \in \vee q_{(\lambda,\mu)} A$. Therefore A is an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal of S.

(2) $\Rightarrow$ (3)

We know that A_α is a subsemiring of S based on Theorem 2. $\forall x \in A_\alpha, \alpha \in (\lambda, \mu]$ and $y \in S$, then $A(x) \geq \alpha$. Since A is a generalized fuzzy weak ideal of S, then $A(xy + yx) \vee \lambda \geq A(x) \wedge \mu \geq \alpha \wedge \mu = \alpha$, by $\lambda < \alpha$, so $A(xy + yx) \geq \alpha$, i.e., $xy + yx \in A_\alpha$.

Therefore A_α is a weak ideal of S, $\forall \alpha \in (\lambda, \mu]$.

(3) $\Rightarrow$ (2)

We obtain that A is a generalized fuzzy complemented subsemiring of S based on Theorem 2. Assume that there exist $x_0, y_0 \in S$ such that $A(x_0y_0 + y_0x_0) \vee \lambda < A(x_0) \wedge \mu$. Choose α such that $A(x_0y_0 + y_0x_0) \vee \lambda < \alpha < A(x_0) \wedge \mu$, then $A(x_0y_0 + y_0x_0) < \alpha$, $A(x_0) > \alpha$ and $\lambda < \alpha < \mu$. So $x_0 \in A_\alpha$. Since A_α is a weak ideal of S, then $A(x_0y_0 + y_0x_0) \geq \alpha$, a contradiction. Therefore A is a generalized fuzzy weak ideal of S.

The above theorem shows that generalized fuzzy weak ideal and $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal are equivalent. Thus we can prove a normal fuzzy set be an $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal by proving it be a generalized fuzzy weak ideal, which is easier than the former. Meanwhile, Theorem 7 establishes a kind of link between generalized fuzzy weak ideal and ordinary weak ideal.

Theorem 8 *Let A and B be generalized fuzzy weak ideal of S, then $A \cap B$ is a generalized fuzzy weak ideal of S.*

Proof We obtain that $A \cap B$ is a generalized fuzzy complemented subsemiring of S based on Theorem 3. For all $x, y \in S$, we have $(A \cap B)(xy + yx) \vee \lambda = (A(xy + yx) \wedge B(xy+yx)) \vee \lambda = (A(xy + yx) \vee \lambda) \wedge (B(xy + yx) \vee \lambda) \geq (A(x) \wedge \mu) \wedge (B(x) \wedge \mu) = (A \cap B)(x) \wedge \mu$.

Therefore $A \cap B$ is a generalized fuzzy weak ideal of S.

Corollary 1 *Let $A_i (i \in I)$ be generalized fuzzy weak ideals of S, then $\cap_{i \in I} A_i$ is a generalized fuzzy weak ideal of S.*

Theorem 9 *Let $A_i (i \in I)$ be generalized fuzzy weak ideals of S, and $\forall i, j \in I, A_i \subseteq A_j$ or $A_j \subseteq A_i$. Then $\cup_{i \in I} A_i$ is a generalized fuzzy weak ideal of S.*

Proof Firstly, we prove that $\vee_{i \in I}(A_i(x) \wedge A_i(y) \wedge \mu) = (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. Obviously, $\vee_{i \in I}(A_i(x) \wedge A_i(y) \wedge \mu) \leq (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. Assume that $\vee_{i \in I} (A_i(x) \wedge A_i(y) \wedge \mu) \neq (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$, then there exists r such that $\vee_{i \in I} (A_i(x) \wedge A_i(y) \wedge \mu) < r < (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. Since $\forall i, j \in I, A_i \subseteq A_j$ or $A_j \subseteq A_i$, then $\exists k \in I$, such that $r < A_k(x) \wedge A_k(y) \wedge \mu$. But $A_i(x) \wedge A_i(y) \wedge \mu < r$, $\forall i \in I$, a contradiction. Thus $\{\vee_{i \in I}(A_i(x) \wedge A_i(y) \wedge \mu)\} = (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. Next, $\forall x, y \in S$, we have $(\cup_{i \in I} A_i)(x + y) \vee \lambda = \vee_{i \in I} A_i(x + y) \vee \lambda = \vee_{i \in I}(A_i(x + y) \vee \lambda) \geq \vee_{i \in I}(A_i(x) \wedge A_i(y) \wedge \mu) = (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. Similarly, we can prove that $(\cup_{i \in I} A_i)(xy) \vee \lambda \geq (\cup_{i \in I} A_i)(x) \wedge (\cup_{i \in I} A_i)(y) \wedge \mu$. $\forall i \in I$, since A_i is a generalized fuzzy complemented subsemiring of S, so $A_i(\overline{a}) \vee \lambda \geq A_i(a) \wedge \mu$. Then $(\vee_{i \in I} A_i(\overline{a})) \vee \lambda \geq A_i(\overline{a}) \vee \lambda \geq A_i(a) \wedge \mu$. We have $(\vee_{i \in I} A_i(\overline{a})) \vee \lambda \geq (\vee_{i \in I} A_i(a)) \wedge \mu$. That is $(\cup_{i \in I} A_i(\overline{a})) \vee \lambda \geq (\cup_{i \in I} A_i(a)) \wedge \mu$.

Thus, $\cup_{i \in I} A_i$ is a generalized fuzzy complemented subsemiring of S.

Finally, $\forall x, y \in S, \cup_{i \in I}(xy + yx) \vee \lambda = \vee_{i \in I} A_i(xy + yx) \vee \lambda = \vee_{i \in I}(A_i(xy + yx) \vee \lambda) \geq \vee_{i \in I}(A_i(x) \wedge \mu) = \vee_{i \in I} A_i(x) \wedge \mu = (\cup_{i \in I} A_i)(x) \wedge \mu$.

Therefore $\cup_{i \in I} A_i$ is a generalized fuzzy weak ideal of S.

Theorem 10 *Let A_1 and A_2 be generalized fuzzy weak ideals of S_1 and S_2 respectively, then $A_1 \times A_2$ is a generalized fuzzy weak ideal of $S_1 \times S_2$.*

Proof Firstly, we prove that $A_1 \times A_2$ is a generalized fuzzy subsemiring of $S_1 \times S_2$. For all $x, y \in S_1 \times S_2$, where $x = (x_1, x_2), y = (y_1, y_2)$, since A_1 and A_2 are generalized fuzzy subsemirings of S_1 and S_2 respectively, then $(A_1 \times A_2)(x + y) \vee \lambda = (A_1 \times A_2)(x_1 + y_1, x_2 + y_2) \vee \lambda = (A_1(x_1 + y_1) \wedge A_2(x_2 + y_2)) \vee \lambda = (A_1(x_1 + y_1) \vee \lambda) \wedge (A_2(x_2 + y_2) \vee \lambda) \geq (A_1(x_1) \wedge A_1(y_1) \wedge \mu) \wedge (A_2(x_2) \wedge A_2(y_2) \wedge \mu) = (A_1 \times A_2)((x_1, x_2)) \wedge (A_1 \times A_2)((y_1, y_2)) \wedge \mu = (A_1 \times A_2)(x) \wedge (A_1 \times A_2)(y) \wedge \mu$.

Similarly, we can prove that $(A_1 \times A_2)(\overline{x}) \vee \lambda \geq (A_1 \times A_2)(x) \wedge \mu$ and $(A_1 \times A_2)(xy) \vee \lambda \geq (A_1 \times A_2)(x) \wedge (A_1 \times A_2)(y) \wedge \mu$. So $A_1 \times A_2$ is a generalized fuzzy complemented subsemiring of $S_1 \times S_2$.

Next, for all $x, y \in S_1 \times S_2$, where $x = (x_1, x_2), y =, (y_1, y_2)$, $(A_1 \times A_2)(xy + yx) \vee \lambda = (A_1 \times A_2)(x_1y_1 + y_1x_1, x_2y_2 + y_2x_2) \vee \lambda = (A_1(x_1y_1 + y_1x_1) \wedge A_2(x_2y_2 + y_2x_2)) \vee \lambda = (A_1(x_1y_1 + y_1x_1) \vee \lambda) \wedge (A_2(x_2y_2 + y_2x_2) \vee \lambda) \geq (A_1(x_1) \wedge \mu) \wedge (A_2(x_2) \wedge \mu) = (A_1 \times A_2)(x_1, x_2) \wedge \mu = (A_1 \times A_2)(x) \wedge \mu$.

Therefore $A_1 \times A_2$ is a generalized fuzzy weak ideal of $S_1 \times S_2$.

Theorem 11 *Let A_i be generalized fuzzy weak ideals of S, then $\prod_{1 \leq i \leq n} A_i$ is a generalized fuzzy weak ideal of $\prod_{1 \leq i \leq n} S_i$.*

Proof Firstly, we prove that $\prod_{1 \leq i \leq n} A_i$ is a generalized fuzzy subsemiring of $\prod_{1 \leq i \leq n} S_i$. For all $x, y \in \prod_{1 \leq i \leq n} S_i$, where $x = (x_1, x_2, \ldots x_n)$ and $y = (y_1, y_2, \ldots y_n)$, then $(\prod_{1 \leq i \leq n} A_i)(x + y) \vee \lambda = \inf A_i(x_i + y_i) \vee \lambda = \inf(A_i(x_i + y_i) \vee \lambda) \geq \inf(A_i(x_i) \wedge A_i(y_i) \wedge \mu) = \inf A_i(x_i) \wedge \inf A_i(y_i) \wedge \mu = (\prod_{1 \leq i \leq n} A_i)(x) \wedge (\prod_{1 \leq i \leq n} A_i)(y) \wedge \mu$. Similarly, we can prove $(\prod_{1 \leq i \leq n} A_i)(xy) \vee \lambda \geq (\prod_{1 \leq i \leq n} A_i)(x) \wedge (\prod_{1 \leq i \leq n} A_i)(y) \wedge \mu$. Next in addition, $(\prod_{1 \leq i \leq n} A_i)(\overline{x}) \vee \lambda = \inf A_i(\overline{x_i}) \vee \lambda = \inf(A_i(\overline{x_i}) \vee \lambda) \geq \inf(A_i(x_i) \wedge \mu) = \inf A_i(x_i) \wedge \mu = (\prod_{1 \leq i \leq n} A_i)(x) \wedge \mu$.

Therefore $\prod_{1 \leq i \leq n} A_i$ is a generalized fuzzy complemented subsimiring of $\prod_{1 \leq i \leq n} S_i$.

Finally, for all $x, y \in \prod_{1 \leq i \leq n} S_i$, where $x = (x_1, x_2, \ldots x_n)$ and $y = (y_1, y_2, \ldots y_n)$, then $(\prod_{1 \leq i \leq n} A_i)(xy + yx) \vee \lambda = \inf A_i(x_iy_i + y_ix_i) \vee \lambda = \inf(A_i(x_iy_i + y_ix_i) \vee \lambda) \geq \inf(A_i(x_i) \wedge \mu) = \inf A_i(x_i) \wedge \mu = (\prod_{1 \leq i \leq n} A_i)(x) \wedge \mu$. Thus $\prod_{1 \leq i \leq n} A_i$ is a generalized fuzzy weak ideal of $\prod_{1 \leq i \leq n} S_i$.

Theorem 12 *Let A be a subset of S, then χ_A is a generalized fuzzy weak ideal of S if and only if A is a weak ideal of S.*

Proof We know that A is a complemented subsemiring of S based on Theorem 4. For all $x \in A$ and $y \in S$, since χ_A is a generalized fuzzy weak ideal of S, then $\chi_A(xy + yx) \vee \lambda \geq \chi_A(x) \wedge \mu = 1 \wedge \mu = \mu$. By $\lambda < \mu$, so $\chi_A(xy + yx) \geq \mu > 0$ and $\chi_A(xy + yx) = 1$. Then $xy + yx \in A$. Therefore A is a weak ideal of S.

Conversely, we can obtain that χ_A is a generalized fuzzy complemented subsemiring of S based on Theorem 4.

Assume that there exist $x_0, y_0 \in S$ such that $\chi_A(x_0y_0 + y_0x_0) \vee \lambda < \chi_A(x_0) \wedge \mu$. Choose α such that $\chi_A(x_0y_0 + y_0x_0) < \alpha < \chi_A(x_0) \wedge \mu < \alpha < \chi_A(x_0) \wedge \mu$, then $\chi_A(x_0y_0 + y_0x_0) < \alpha, \chi_A(x_0) > \alpha$ and $\lambda < \alpha < \mu$. i.e. $x_0 \in A$. Since A is a weak ideal of S then $x_0y_0 + y_0x_0 \in A$. So $\chi_A(x_0y_0 + y_0x_0) = 1 > \alpha$, a contradiction.

Thus χ_A is a generalized fuzzy weak ideal of S.

Theorem 13 *Let A be a generalized fuzzy weak ideals of S, then* $A_{\underline{\lambda}} = \{x|A(x) > \lambda\}$ *is a generalized weak ideal of S.*

Proof Firstly, we prove that $A_{\underline{\lambda}}$ is a complemented subsemiring of S. For all $x, y \in A_{\underline{\lambda}}$, since A is a generalized fuzzy weak ideals of S, then $A(x+y) \vee \lambda \geq A(x) \wedge A(y) \wedge \mu > \lambda$, so $A(x+y) > \lambda$, i.e., $x + y \in A_{\underline{\lambda}}$. Similarly, we can prove that $xy, \bar{x} \in A_{\underline{\lambda}}$.

Therefore $A_{\underline{\lambda}}$ is a complemented submiring of S. Next, for all $x \in A_{\underline{\lambda}}$ and $y \in S$, then $A(x) > \lambda$. Since A is a generalized fuzzy weak ideals of S, then $A(xy + yx) \vee \lambda \geq A(x) \wedge \mu > \lambda$, so $A(xy + yx) > \lambda$, i.e., $xy + yx \in A_{\underline{\lambda}}$.

Therefore $A_{\underline{\lambda}}$ is a weak ideal of S.

Theorem 14 *Let* $f : S \to H$ *be a full homomorphism, if A is a generalized fuzzy weak ideal of S, then* $f(A)$ *is a generalized fuzzy weak ideal of H.*

Proof Based on Theorem 5, we know that $f(A)$ is a generalized fuzzy complemented subsemiring of H. For all $z_1, z_2 \in H$, there exist $x_1, x_2 \in S$, such that $f(x_1) = z_1, f(x_2) = z_2$, then $f(x_1x_2 + x_2x_1) = f(x_1x_2) + f(x_2x_1) = f(x_1)f(x_2) + f(x_2)f(x_1) = z_1z_2 + z_2z_1$, So $f(A)(z_1z_2 + z_2z_1) \vee \lambda = \sup\{A(x)|f(x) = z_1z_2 + z_2z_1\} \vee \lambda = \sup\{A(x) \vee \lambda|f(x) = z_1z_2 + z_2z_1\} \geq \sup\{A(x_1x_2 + x_2x_1) \vee \lambda|f(x_1) = z_1, f(x_2) = z_2\} \geq \sup\{A(x_1) \wedge \mu|f(x_1) = z_1\} = \sup\{A(x_1) \wedge \mu|f(x_1) = z_1\} == f(A)(z_1) \wedge \mu$.

Therefore $f(A)$ is a generalized fuzzy ideal of H.

Theorem 15 *Let* $f : S \to H$ *be a homomorphism, if B is a generalized fuzzy weak ideal of H, then* $f^{-1}(B)$ *is a generalized fuzzy weak ideal of S.*

Proof Based on Theorem 6, we can obtain that $f^{-1}(B)$ is a generalized fuzzy complemented subsemiring of S. For all $x, y \in S$, then $f(x), f(y) \in H$. Since B is a generalized fuzzy weak ideal of H, so $f^{-1}(B)(xy + yx) \vee \lambda = B(f(xy+yx)) \vee \lambda = B(f(x)f(y) + f(y)f(x)) \vee \lambda \geq B(f(x)) \wedge \mu = f^{-1}(B)(x) \wedge \mu$.

Therefore $f^{-1}(B)$ is a generalized fuzzy weak ideal of S.

4 Conclusion

In this present investigation, the concepts of generalized fuzzy weak ideal and $(\in, \in \vee q_{(\lambda,\mu)})$-fuzzy weak ideal are proposed. Moreover, the relevant properties are studied. Further we will do some relevant properties on chain condition of fuzzy weak ideal.

Acknowledgments This work is supported by Program for Innovative Research Team of Jiangnan University(No:200902).

References

1. Rosenfeld, A.: Fuzzy groups. J. Math. Anal. Appl. **35**, 512–517 (1971)
2. Liao, Z.H., Chen, M.: Fuzzy subsemigroups and fuzzy completely regular subsemigroups. J. Jiangnan Univ. (Natural Science) **8**, 242–244 (2009)
3. Liao, Z.H., Hu, M.H., Chen, M., et al.: Generalized fuzzy interior ideals in semigroups. In: Third International Joint Conference on Computational Science and Optimization, pp. 197–200 (2010)
4. Vandiver, H.S.: On some simple types of semirings. Am. Math. Monthly **46**, 22–26 (1939)
5. Glazek, K.: A Guide to the Literature on Semirings and their Applications in Mathematics and Information Sciences. Kluwer, Dordrecht (2002)
6. Golan, J.S.: Semirings and Affine Equations Over Them: Theory and Applications. Kluwer, Dordrecht (2003)
7. Liu, W.J.: Fuzzy invariant subgroups and fuzzy ideals. Fuzzy Sets Syst. **8**, 133–139 (1982)
8. Mukherjee, T.K., Sen, M.K.: On fuzzy ideals of a ring. Fuzzy Sets Syst. **21**, 99–104 (1987)
9. Mukherjee, T.K., Sen, M.K.: Prime fuzzy ideals in rings. Fuzzy Sets Syst. **32**, 337–341 (1989)
10. Feng, J., Zhan, C.: Complemented semirings congruence. Fuzzy Syst. Math. **24**, 61–65 (2010)
11. Ming, P.P., Ming, L.Y.: Fuzzy topology: neighbourhood structure of a fuzzy point and Moore-Smith convergence. J. Math. Anal. Appl. **76**, 571–599 (1980)
12. Zhu, C., Liao, Z.H., Luo, X.T., et al.: Fuzzy complemented sumirings. Fuzzy Syst. Math. **27**(5), 47–54 (2013)
13. Liao, Z.H., Gu, H.: Fuzzy normal subgroup. Fuzzy Syst. Math. **20**, 340–343 (2006)
14. Yao, B.X.: Fuzzy Theory on Groups and Rings, pp. 1–125. Science Press, Beijing (2008)

Generalized Fuzzy Sets and Fuzzy Relations

Yan-cai Zhao, Zu-hua Liao, Teng Lu and Juan Tong

Abstract In classical fuzzy set theory, a fuzzy set is a membership function which associates with each element a real number in $[0, 1]$, a fuzzy relation is a function which associates with each pair of elements a real number in $[0, 1]$. In the present paper, we generalize the above two concepts by associating with each set a real number in $[0, 1]$, and associating with each pair of sets a real number in $[0, 1]$, respectively. We then give a series of properties for these two types of generalized concepts. We also show that a generalized fuzzy relation can be induced by a classical fuzzy relation, which shows the communication of our generalized fuzzy concepts with the classical fuzzy theory.

Keywords Fuzzy set · Fuzzy relation · Generalized fuzzy set · Generalized fuzzy relation · Power set

1 Introduction

The concept of a fuzzy set was introduced by Zadeh [10]. A *fuzzy set* in a referential (universe of discourse) X is characterized by a membership function A which associates with each element $x \in X$ a real number $A(x) \in [0, 1]$, having the interpretation that $A(x)$ is the membership degree of x in the fuzzy set A. For convenience, we also call the set X in above definition the *base set* of a the fuzzy set A.

Let X, Y be two sets. A mapping $R : X \times Y \to [0, 1]$ is called a *fuzzy relation* [10]. The number $R(x, y) \in [0, 1]$ can be interpreted as the degree of relationship between x and y.

Y. Zhao (✉)
Wuxi City College of Vocational Technology, 214153 Jiangsu, China
e-mail: zhaoyc69@126.com

Y. Zhao · Z. Liao (✉) · T. Lu · J. Tong
School of Science, Jiangnan University, 214122 Jiangsu, China
e-mail: liaozuhua57@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_2

Since fuzzy set was introduced by Zadeh in 1965 [10], many extensions have been developed, such as intutionistic fuzzy set [1], type-2 fuzzy set [2, 4], type-*n* fuzzy set [2], fuzzy multiset [5, 9] and hesitant fuzzy set [6–8, 11]. So far as we know, all types of fuzzy set assign a value or a set of values to an element of a classical set. However, there are many situations in our real lives in which one has to make decision on a whole set. For example, a patient usually has several symptoms. His/Her doctor has to make a decision by whole of his/her symptoms, which means that the doctor assigns a value to the set of all the patient's symptoms. We will give the details of this example later.

In this paper, we first generalize the classical fuzzy set on a base set X to one on a collection of sets. Then, we concentrate on the study of the generalized fuzzy sets on base set $\mathcal{P}(X)$, the power set of X, and obtain a series of properties; We also generalize the classical fuzzy relation between two elements to several types of relations between two sets, and obtain a series of properties for these generalized fuzzy relations. Among them, we construct a equivalence fuzzy relation at the end of the paper.

2 Generalized Fuzzy Sets and Their Operations

In our real lives, there exist many situations in which we should make decisions on a collection of sets.

Example 1 A doctor usually judge that if a patient has caught a cold by the following symptoms of this patient: Fever (F for short), Headache (H for short) and Cough (C for short). The following table give the corresponding numbers to different sets of symptoms. Each number means a degree of a patient catch a cold.

In the following table, the collection D of decisions is a fuzzy set on the collection of the set of different symptoms of a patient (Fig. 1).

$$D = \frac{0}{\emptyset} + \frac{0.2}{\{F\}} + \frac{0.1}{\{H\}} + \frac{0.2}{\{C\}} + \frac{0.3}{\{F,H\}} + \frac{0.5}{\{F,C\}} + \frac{0.4}{\{H,C\}} + \frac{0.9}{\{F,H,C\}}.$$

A classical fuzzy set assigns a value to an element. We now give a generalized type of fuzzy set which assigns a value to a set.

Definition 1 Let $\mathcal{S} = \{S_i |\ i \in I\}$ be a collection of sets. Then a mapping A such that $A : \mathcal{S} \to [0, 1], S \mapsto A(S), \forall S \in \mathcal{S}$ is called a fuzzy set on $\mathcal{S}$. $A(S)$ is called the *membership degree* of S in $\mathcal{S}$. Let the collection of all the fuzzy sets on $\mathcal{S}$ be $\mathcal{F}(\mathcal{S})$.

Fig. 1 Fuzzy decisions on the power set of a set

symptoms	$\emptyset$	$\{F\}$	$\{H\}$	$\{C\}$	$\{F, H\}$	$\{F, C\}$	$\{H, C\}$	$\{F, H, C\}$
decision	0	0.2	0.1	0.2	0.3	0.5	0.4	0.9

As an special collection of sets, the power set of a set is interesting. So we will concentrate our studies on the fuzzy sets on a power set.

Definition 2 Let X be a set. Then a mapping $A_{\mathcal{P}}$ such that

$$A_{\mathcal{P}} : \mathcal{P}(X) \to [0, 1], P \mapsto A_{\mathcal{P}}(P), \forall P \in \mathcal{P}(X)$$

is called a fuzzy set on $\mathcal{P}(X)$. $A_{\mathcal{P}}(P)$ is called the *membership degree* of P in $\mathcal{P}(X)$. Let the collection of all the fuzzy sets on $\mathcal{P}(X)$ be $\mathcal{F}(\mathcal{P}(X))$.

Give a fuzzy set on X, we can deduce a fuzzy set on $\mathcal{P}(X)$ as follows.

Definition 3 Given a fuzzy set A on X, the mapping such that

$$A_{max} : \mathcal{P}(X) \to [0, 1], P \mapsto A_{max}(P) = \vee_{x\in P} A(x)$$

is called the *max-type induced fuzzy set* by A. Denote $\mathcal{F}_{max}(\mathcal{P}(X)) = \{A_{max} \mid A \in \mathcal{F}(X)\}$.

Definition 4 Given a fuzzy set A on X, the mapping such that

$$A_{min} : \mathcal{P}(X) \to [0, 1], P \mapsto A_{min}(P) = \vee_{x\in P} A(x)$$

is called the *min-type induced fuzzy set* by A. Denote $\mathcal{F}_{min}(\mathcal{P}(X)) = \{A_{min} \mid A \in \mathcal{F}(X)\}$.

Given any fuzzy set A on X, then A induces a relation $\sim$ on X by defining that $x \sim y \Leftrightarrow A(x) = A(y)$. It is easy to see that $\sim$ is an equivalence relation on X. Therefore, there exists the quotient set $X/\sim$. Similarly, A_{max} or A_{min} induces a relation $\sim'$ on $\mathcal{P}(X)$ such that $P_1 \sim' P_2 \Leftrightarrow A_{\mathcal{P}}(x) = A_{\mathcal{P}}(y)$ for any $P_1, P_2 \in \mathcal{P}(X)$. $\sim'$ is an equivalence relation on $\mathcal{P}(X)$ and thus there exists the quotient set $\mathcal{P}(X)/\sim'$.

The order relation of elements in a set is usually defined as $x \leq y \Leftrightarrow A(x) \leq A(y)$. Now we define the order relation of two sets as follows.

Definition 5 Let A be a fuzzy set on X. For any $P_1, P_2 \in \mathcal{P}(X), P_1 \leq P_2 \Leftrightarrow A_{\mathcal{P}}(P_1) \leq A_{\mathcal{P}}(P_2)$.

We further give a more general order relation on $\mathcal{P}(X)$ as follows.

Definition 6 Let $\widetilde{A}, \widetilde{B} \in \mathcal{F}(\mathcal{P}(X))$. If $\forall P \in \mathcal{P}(X), \widetilde{A}(P) \leq \widetilde{B}(P)$, then we say $\widetilde{A} \subseteq \widetilde{B}$. If $\forall P \in \mathcal{P}(X), \widetilde{A}(P) = \widetilde{B}(P)$, then we say that $\widetilde{A} = \widetilde{B}$.

Theorem 1 *Let* $\widetilde{A}, \widetilde{B}, \widetilde{C} \in \mathcal{F}(\mathcal{P}(X))$*, then we have the follows.*

(1) Self-reflexivity. $\widetilde{A} \subseteq \widetilde{A}$.
(2) Anti-symmetry. $\widetilde{A} \subseteq \widetilde{B}, \widetilde{B} \subseteq \widetilde{A} \Rightarrow \widetilde{A} = \widetilde{B}$.
(3) Transitivity. $\widetilde{A} \subseteq \widetilde{B}, \widetilde{B} \subseteq \widetilde{C} \Rightarrow \widetilde{A} \subseteq \widetilde{C}$.

From Theorem 1 we know that $(\mathcal{F}(\mathcal{P}(X)), \subseteq)$ is a partially ordered set.

3 Generalized Fuzzy Relations

The classical fuzzy set theory defined the relations between two elements as follows. Let X, Y be two classical sets. A mapping $R : X \times Y \to [0, 1]$ is called a *fuzzy relation* [10]. The number $R(x, y) \in [0, 1]$ can be interpreted as the degree of relationship between x and y.

Now we define the fuzzy relation between two sets as follows.

Definition 7 Let X, Y be two classical sets. A mapping $R : \mathcal{P}(X) \times \mathcal{P}(Y) \to [0, 1]$ is called a *fuzzy relation*. The number $R(A, B) \in [0, 1]$ can be interpreted as the degree of relationship between A and B.

Further, we give two types of induced relations between two sets as follows.

Definition 8 Let $R : X \times X \to [0, 1]$ be a fuzzy relation on X. The mapping $\widetilde{R} : \mathcal{P}(X) \times \mathcal{P}(X) \to [0, 1]$ such that $\widetilde{R}(A, B) = \bigvee_{a \in A, b \in B} R(a, b)$ for $A, B \in \mathcal{P}(X)$ is called *the max-type induced fuzzy relation between A and B.*

Definition 9 Let $R : X \times X \to [0, 1]$ be a fuzzy relation on X. The mapping $\widetilde{R} : \mathcal{P}(X) \times \mathcal{P}(X) \to [0, 1]$ such that $\widetilde{R}(A, B) = \bigwedge_{a \in A, b \in B} R(a, b)$ for $A, B \in \mathcal{P}(X)$ is called *the min-type induced fuzzy relation between A and B.*

Theorem 2 *If R is a fuzzy relation on X, then the max-type induced fuzzy relation $\widetilde{R}$ on $\mathcal{P}(X)$ has the following properties.*

(1) $A \subseteq C, B \subseteq D \Rightarrow \widetilde{R}(A, B) \leq \widetilde{R}(C, D)$;
(2) If R is self-reflexive and $A \cap B \neq \emptyset$, *then* $\widetilde{R}(A, B) = 1$;
(3) $\widetilde{R}(A, B \cup C) = \widetilde{R}(A, B) \vee \widetilde{R}(A, C)$;
(4) $\widetilde{R}(A, B \cap C) \leq \widetilde{R}(A, B) \wedge \widetilde{R}(A, C)$.

Proof (1). By Definition 8, it is easy to see.
(2). Since R is self-reflexive, $R(x, x) = 1$ for any $x \in X$. Choose an element $Y \in A \cap B$. Then by (1) $\widetilde{R}(A, B) \geq \widetilde{R}(\{x\}, \{x\}) = R(x, x) = 1$. So $\widetilde{R}(A, B) = 1$.
(3).

$$\begin{aligned}
\widetilde{R}(A, B \cup C) &= \bigvee_{\substack{a \in A \\ u \in B \cup C}} R(a, u) \\
&= \bigvee_{\substack{a \in A \\ u \in B}\ or\ \substack{a \in A \\ u \in C}} R(a, u) \\
&= \left\{ \bigvee_{\substack{a \in A \\ u \in B}} R(a, u) \right\} \vee \left\{ \bigvee_{\substack{a \in A \\ u \in C}} R(a, u) \right\} \\
&= \widetilde{R}(A, B) \vee \widetilde{R}(A, C).
\end{aligned}$$

(4). It is easy to be deduced from (1).

Theorem 3 *If R is a fuzzy relation on X, then the min-type induced fuzzy relation $\widetilde{R}$ on $\mathcal{P}(X)$ has the following properties.*

(1) $A \subseteq C, B \subseteq D \Rightarrow \widetilde{R}(A, B) \geq \widetilde{R}(C, D)$;
(2) *If R is self-reflexive and $A - B \neq \emptyset$ (Res. $B - A \neq \emptyset$), then* $\widetilde{R}(A, B) = \widetilde{R}(A - B, B)$ *(Res.* $\widetilde{R}(A, B) = \widetilde{R}(A, B - A)$*)*;
(3) $\widetilde{R}(A, B \cup C) = \widetilde{R}(A, B) \wedge \widetilde{R}(A, C)$;
(4) $\widetilde{R}(A, B \cap C) \geq \widetilde{R}(A, B) \vee \widetilde{R}(A, C)$.

Proof The proofs are similar to those in Theorem 2, and thus we omit.

Now we give two types of definitions of transitivity of fuzzy relations on $\mathcal{P}(X)$, and will provide an equivalence fuzzy relation on $\mathcal{P}(X)$.

Definition 10 A fuzzy relation R on $\mathcal{P}(X)$ is called *I-type transitive*, if $R^2 = R \circ R \subseteq R$, that is, $\forall (U, V) \in \mathcal{P}(X) \times \mathcal{P}(X)$, $\bigvee_{W \in \mathcal{P}(X)} R(U, W) \bigwedge R(W, V) \leq R(U, V)$.

Definition 11 A fuzzy relation R on $\mathcal{P}(X)$ is called *II-type transitive*, if $R^2 = R \circ R \subseteq R$, that is, $\forall (U, V) \in \mathcal{P}(X) \times \mathcal{P}(X)$, $\bigvee_{\{x\} \in \mathcal{P}(X)} R(U, \{x\}) \bigwedge R(\{x\}, V) \leq R(U, V)$.

Definition 12 Let $R : X \times X \to [0, 1]$ be a fuzzy relation on X. For any two different elements $x, y \in X$, a path from x to y, denoted $P(x, y)$, is a set of continuous pairs $(x, a_1), (a_1, a_2), (a_2, a_3), \ldots, (a_n, y)$, where, each pair is called an *edge* of the path. Let $E(P(x, y))$ be the set of all edges in a path $P(x, y)$. The *degree* of a path $P(x, y)$ is $S(P(x, y)) = \bigwedge_{e \in E(P(x,y))} R(e)$. Suppose that there are l paths $P_1, P_2, \ldots, P_l$ from x to y. Then the *connective degree* between x and y is

$$S(x, y) == \begin{cases} \bigvee_{i=1}^{l} S(P_i), & \text{if } x \neq y; \\ 1, & \text{if } x = y. \end{cases}$$

The set of paths from A to B is $P(A, B) = \{P(a, b) | a \in A, b \in B\}$. The *connective degree* between A and B is $S(A, B) = \bigvee_{a \in A, b \in B} S(a, b)$.

The induced fuzzy relation S on $\mathcal{P}(X)$ has the following properties.

Theorem 4 *If R is a fuzzy relation on X, then the fuzzy relation S on $\mathcal{P}(X)$ has the following properties.*

(1) $A \subseteq C, B \subseteq D \Rightarrow S(A, B) \leq S(C, D)$;
(2) *If* $A \cap B \neq \emptyset$, *then* $S(A, B) = 1$;
(3) $S(A, B \cup C) = S(A, B) \vee S(A, C)$;
(4) $S(A, B \cap C) \leq S(A, B) \wedge S(A, C)$.

Proof The proof is similar to that of Theorem 2, and thus we omit.

Theorem 5 *If R is a symmetric fuzzy relation on X, then S is an equivalence relation on $\mathcal{P}(X)$, under the meaning of the II-type transitivity.*

Proof By (2) of Theorem 4, $S(A, A) = 1$, and thus S is self-reflexive. Since R is symmetric, it is easy to see that S is symmetric. It remains to prove that S satisfies transitivity. Note that

$$S^2(A, B) = \bigvee_{w \in X} (S(A, \{w\}) \wedge S(\{w\}, B)), \forall A, B \in \mathcal{P}(X).$$

If $A \cap B \neq \emptyset$, then $S(A, B) = 1$, we have $S(A, B) \geq S^2(A, B)$. Now assume that $A \cap B = \emptyset$. Choose an arbitrary element $w \in X$. If $w \in A$, then $S(A, \{w\}) = 1$. It follows that $S(A, B) \geq S(\{w\}, B) = S(A, \{w\}) \wedge S(\{w\}, B)$. If $w \in B$, then $S(\{w\}, B) = 1$. It follows that $S(A, B) \geq S(A, \{w\}) = S(A, \{w\}) \wedge S(\{w\}, B)$. When $w \notin A \cup B$, let P_1 and P_2 are the optimal paths of $(A, \{w\})$ and $(\{w\}, B)$, respectively. Then $P_1 \cup P_2$ contains a path P of (A, B), and thus $E(P) \subseteq E(P_1) \cup E(P_2)$. Therefore, $S(A, B) \geq S(P) \geq S(P_1) \wedge S(P_2) = S(A, \{w\}) \wedge S(\{w\}, B)$. In either case, $S(A, B) \geq S(A, \{w\}) \wedge S(\{w\}, B)$. So $S(A, B) \geq S^2(A, B)$, which means the transitivity of $S(A, B)$.

Acknowledgments This article is completed during the first author being a visiting scholar in Jiangnan University. Research was partially supported by the domestic senior visiting scholar program in higher occupation colleges in Jiangsu province (2014FX075).

The authors are grateful to the referees for their valuable suggestions, which result in the present version of the paper. This paper was also supported by the Natural Science Foundation of Jiangsu Province (No. BK20151117).

References

1. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**, 87–96 (1986)
2. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
3. Hu, B.Q.: Foudation of Fuzzy Theory. Wuhan University Press (2010)
4. Miyamoto, S.: Remarks on basics of fuzzy sets and fuzzy multisets. Fuzzy Sets Syst. **156**, 427–431 (2005)
5. Miyamoto, S.: Multisets and fuzzy multisets. In: Liu, Z.Q., Miyamoto, S. (eds.) Soft Computing and Human-Centered Machines, pp. 9–33. Springer, Berlin (2000)
6. Torra, V., Narukawa, Y.: On hesitant fuzzy sets and decision. In: The 18th IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, pp. 1378–1382 (2009)
7. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**, 529–539 (2010)
8. Xu, Z.: Hesitant fuzzy sets theory. Stud. Fuzziness Soft Comput. **314**(1) (2014). (Springer International Publishing Switzerland). doi:10.1007/978-3-319-04711-91
9. Yager, R.R.: On the theory of bags. Int. J. Gen. Syst. **13**, 23–37 (1986)
10. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
11. Zhu, B., Xu, Z.S., Xia, M.M.: Dual hesitant fuzzy sets. J. Appl. Math. **13** (2012). doi:10.1155/2012/879629

Fuzzy Topologies and Fuzzy Preorders Based on Complete Co-residuated Lattices

Piwei Chen and Yu Zeng

Abstract For a complete co-residuated lattice L, the concepts of $\overline{L}$-topological spaces and $\overline{L}$-preorders are introduced by virtue of the nonmembership functions in the intuitionistic fuzzy set theory. Furthermore, the methods to induce one from another are studied in order to establish a stable relationship between each other.

Keywords Fuzzy topology · Fuzzy order · Complete co-residuated lattice · Many-valued logic

1 Introduction

Since Zadeh [16] proposed fuzzy sets and Chang [2] introduced fuzzy theory into topology, many authors discussed various aspects of fuzzy topology [11, 13, 15, 17]. Fuzzy topology theory has formed its own research pattern. On the other hand, fuzzy order is a basic concept in fuzzy set theory and plays an important role in many branches of fuzzy set theory, which is closely related to many basic concepts in mathematics [1]. From a mathematical viewpoint the important feature of fuzzy set theory is the replacement of the two-valued logic by a many-valued logic [8]. In the setting of complete residuated lattice being the truth table of many-valued logic, many authors discussed the close relationships between fuzzy topologies and fuzzy orders [3–5, 10, 14].

This paper is devoted to a special topic in the research of the interrelationship between fuzzy topologies and fuzzy orders in the setting of complete co-residuated lattice being the truth table.

P. Chen
School of Mathematical Sciences, Ocean University of China, Qingdao 266100, China
e-mail: cpw@ouc.edu.cn

Y. Zeng (✉)
College of Basic Education, Qingdao University of Technology, Qingdao 266300, China
e-mail: zhugewochong@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_3

2 Preliminaries

In this section, we recall some basic definitions and results about lattices and fuzzy sets, which are needed in this paper. Readers can refer to [6, 7, 9, 12].

A complete co-residuated lattice is a triad $(L, \oplus, \ominus)$, where L is a complete lattice with top element 1 and bottom element 0, and $\oplus, \ominus : L \times L \longrightarrow L$ are binary operations on L such that

(i) $\oplus$ is monotone on each variable;
(ii) $(L, \oplus, 0)$ is a commutative monoid;
(iii) $a \leq b \oplus c \iff a \ominus b \leq c$ for all $a, b, c \in L$.

Example 1 Let $L = [0, 1]$ be the unit interval, $a \oplus b = a \vee b$,

$$a \ominus b = \begin{cases} 0, \ a \leq b; \\ a, \ a > b. \end{cases}$$

Then $([0, 1], \vee, \ominus)$ is a complete co-residuated lattice.

In this paper, $(L, \oplus, \ominus)$ is always assumed to be a complete co-residuated lattice. We also write simply L for $(L, \oplus, \ominus)$ when there will be no confusion. Some basic properties of complete co-residuated lattices are collected in the following proposition.

Proposition 1 *Suppose* $(L, \oplus, \ominus)$ *is a complete co-residuated lattice. Then*

(I1) $1 \oplus a = 1;\ \ 0 \oplus a = a;$
(I2) $a \ominus 0 = a;\ \ a \ominus 1 = 0;\ \ a \ominus a = 0;$
(I3) $a \ominus c \leq (a \ominus b) \oplus (b \ominus c);$
(I4) $a \oplus \bigwedge_{j \in J} b_j = \bigwedge_{j \in J}(a \oplus b_j);$
(I5) $(\bigvee_{j \in J} a_j) \ominus b = \bigvee_{j \in J}(a_j \ominus b);$
(I6) $a \ominus \bigwedge_{j \in J} b_j = \bigvee_{j \in J}(a \ominus b_j).$

Let X be a nonempty set and L a complete lattice. An intuitionistic L-fuzzy set A on X is an object having the form

$$A = \{\langle x, \underline{A}(x), \overline{A}(x)\rangle : x \in X\}$$

where the function $\underline{A} : X \to L$ and $\overline{A} : X \to L$ denote the degree of membership (namely $\underline{A}(x)$) and the degree of nonmembership (namely $\overline{A}(x)$) of each element $x \in X$ to the set A. In this sense, we call the object $A = \{\langle x, \underline{A}(x)\rangle : x \in X\}$ an L-subset of X, and $A = \{\langle x, \overline{A}(x)\rangle : x \in X\}$ an $\overline{L}$-subset of X respectively.

Definition 1 Let L be a complete co-residuated lattice. $A = \{\langle x, \overline{A}(x)\rangle : x \in X\}$ and $B = \{\langle x, \overline{B}(x)\rangle : x \in X\}$ are two $\overline{L}$-subsets of X. For any $a \in L$, define

(1) $A \subset B$ iff $\overline{A}(x) \geq \overline{B}(x)$ for all $x \in X$;
(2) $A = B$ iff $A \subset B$ and $B \subset A$;
(3) $A \cup B = \{\langle x, \overline{A \cup B}(x)\rangle : x \in X\}$ where $\overline{A \cup B}(x) = \overline{A}(x) \wedge \overline{B}(x)$;
(4) $A \cap B = \{\langle x, \overline{A \cap B}(x)\rangle : x \in X\}$ where $\overline{A \cap B}(x) = \overline{A}(x) \vee \overline{B}(x)$;
(5) $a_X = \{\langle x, \overline{a_X}(x)\rangle : x \in X\}$ where $\overline{a_X}(x) = a$ for all $x \in X$;
(6) $A \oplus a = \{\langle x, \overline{A \oplus a}(x)\rangle : x \in X\}$ where $\overline{A \oplus a}(x) = \overline{A}(x) \oplus a$;
(7) $A \ominus a = \{\langle x, \overline{A \ominus a}(x)\rangle : x \in X\}$ where $\overline{A \ominus a}(x) = \overline{A}(x) \ominus a$.

3 Fuzzy Topologies and Fuzzy Orders Valued by a Complete Co-residuated Lattice

Definition 2 Let X be a nonempty set and L a complete co-residuated lattice. If the family $\mathcal{T}$ of $\overline{L}$-subsets of X satisfies

(O1) $1_X \in \mathcal{T},\ 0_X \in \mathcal{T}$;
(O2) For any $A, B \in \mathcal{T}$, $A \cap B \in \mathcal{T}$;
(O3) For any $\{A_j \mid j \in J\} \subset \mathcal{T}$, $\bigcup_{j \in J} A_j \in \mathcal{T}$,

then $\mathcal{T}$ is called an $\overline{L}$-topology on X. Further, if an $\overline{L}$-topology $\mathcal{T}$ on X satisfies

(O4) For any $a \in L$ and $A \in \mathcal{T}$, $A \oplus a \in \mathcal{T}$;
(O5) For any $a \in L$ and $A \in \mathcal{T}$, $A \ominus a \in \mathcal{T}$,

then $\mathcal{T}$ is called a strong $\overline{L}$-topology on X.

Definition 3 An $\overline{L}$-topology $\mathcal{T}$ on X is Alexandroff if

(O2') For any $\{A_j \mid j \in J\} \subset \mathcal{T}$, $\bigcap_{j \in J} A_j \in \mathcal{T}$,

Example 2 Let $X = \{x, y\}$ and $L = ([0, 1], \vee, \ominus)$. Then

$$\mathcal{T} = \{A \mid \overline{A}(x) < 1,\ \overline{A}(y) < 1\} \cup \{1_X\}$$

is a strong $\overline{L}$-topology on X. However, $(X, \mathcal{T})$ is not Alexandroff.

In fact, for any $a \in [0, 1]$, put an $\overline{L}$-subset A_a of $X = \{x, y\}$ as $\overline{A_a}(x) = a$ and $\overline{A_a}(y) = 0$. Then for any $a \in [0, 1)$ we have $A_a \in \mathcal{T}$. But

$$\bigcap_{a \in [0,1)} A_a = A_1 \notin \mathcal{T}.$$

Theorem 1 *Let L be a complete co-residuated lattice and $\mathcal{T}$ an $\overline{L}$-topology on X. Then for any $\overline{L}$-subset A of X, define the interior of A as*

$$int(A) = \bigcup\{G \mid G \in \mathcal{T}, G \subset A\}.$$

Then

(I1) $int(0_X) = 0_X$;
(I2) $int(A) \subset A$;
(I3) $int(A \cap B) = int(A) \cap int(B)$;
(I4) $int(int(A)) = int(A)$;

In addition, if $\mathcal{T}$ is a strong $\overline{L}$-topology on X, then

(I5) $int(A \ominus a) = int(A) \ominus a$ *for any* $a \in L$.

Proof The proofs of (I1), (I2), (I3) and (I4) are trivial. Now suppose that $\mathcal{T}$ is a strong $\overline{L}$-topology on X and $a \in L$.

Since $int(A) \in \mathcal{T}$, we have $int(A) \ominus a \in \mathcal{T}$. Then $int(A) \ominus a \subset A \ominus a$, Therefore $int(A) \ominus a \subset int(A \ominus a)$.

On the other hand, for any $x \in X$,

$$\begin{aligned}
\overline{int(A \ominus a)}(x) \oplus a &= \overline{\bigcup\{G \mid G \in \mathcal{T}, G \subset A \ominus a\}}(x) \oplus a \\
&= \bigwedge\{\overline{G}(x) \mid G \in \mathcal{T}, G \subset A \ominus a\} \oplus a \\
&= \bigwedge\{\overline{G}(x) \oplus a \mid G \in \mathcal{T}, G \subset A \ominus a\} \\
&= \bigwedge\{\overline{(G \oplus a)}(x) \mid G \in \mathcal{T}, G \oplus a \subset A\} \\
&\geq \bigwedge\{\overline{H}(x) \mid H \in \mathcal{T}, H \subset A\} \\
&= \overline{\bigcup\{H \mid H \in \mathcal{T}, H \subset A\}}(x) \\
&= \overline{int(A)}(x).
\end{aligned}$$

So $\overline{int(A)}(x) \ominus a \leq \overline{int(A \ominus a)}(x)$. Therefore $int(A \ominus a) \subset int(A) \ominus a$.

Definition 4 Let X be a nonempty set and L a complete co-residuated lattice. If the $\overline{L}$-subset R of $X \times X$ satisfies

(P1) $\overline{R}(x, x) = 0$ for all $x \in X$;
(P2) $\overline{R}(x, z) \leq \overline{R}(x, y) \oplus \overline{R}(y, z)$ for all $x, y, z \in X$,

then R is called an $\overline{L}$-preorder on X.

Theorem 2 *For any $\overline{L}$-preorder R on X, set a family of $\overline{L}$-subsets of X as follows:*

$$\Gamma(R) = \{\mu \mid \overline{\mu}(y) \ominus \overline{\mu}(x) \leq \overline{R}(x, y) \textit{ for any } x, y \in X\}.$$

Then $\Gamma(R)$ is an Alexandroff strong $\overline{L}$-topology on X.

Proof (O1): For any $x, y \in X$, $\overline{1_X}(y) \ominus \overline{1_X}(x) = 1 \ominus 1 = 0 \leq \overline{R}(x, y)$. Then $1_X \in \Gamma(R)$.

For any $x, y \in X$, $\overline{0_X}(y) \ominus \overline{0_X}(x) = 0 \ominus 0 = 0 \leq \overline{R}(x, y)$. Then $0_X \in \Gamma(R)$.

(O2') and (O3): Suppose $\{A_j \mid j \in J\} \subset \Gamma(R)$. For any $x, y \in X$,

$$\overline{A_j}(y) \leq \overline{A_j}(x) \oplus \overline{R}(x, y) \leq \bigvee_{j \in J} \overline{A_j}(x) \oplus \overline{R}(x, y).$$

Then

$$\bigvee_{j \in J} \overline{A_j}(y) \leq \bigvee_{j \in J} \overline{A_j}(x) \oplus \overline{R}(x, y).$$

That is,

$$\overline{\bigcap_{j \in J} A_j}(y) \leq \overline{\bigcap_{j \in J} A_j}(x) \oplus \overline{R}(x, y).$$

So

$$\overline{\bigcap_{j \in J} A_j}(y) \ominus \overline{\bigcap_{j \in J} A_j}(x) \leq \overline{R}(x, y).$$

We have $\bigcap_{j \in J} A_j \in \Gamma(R)$.

$$\begin{aligned} \overline{\bigcup_{j \in J} A_j}(x) \oplus \overline{R}(x, y) &= \bigwedge_{j \in J} \overline{A_j}(x) \oplus \overline{R}(x, y) = \bigwedge_{j \in J} (\overline{A_j}(x) \oplus \overline{R}(x, y)) \\ &\geq \bigwedge_{j \in J} \overline{A_j}(y) = \overline{\bigcup_{j \in J} A_j}(y). \end{aligned}$$

So

$$\overline{\bigcup_{j \in J} A_j}(y) \ominus \overline{\bigcup_{j \in J} A_j}(x) \leq \overline{R}(x, y).$$

Therefore, $\bigcup_{j \in J} A_j \in \Gamma(R)$.

(O4) and (O5): Suppose $a \in L$ and $A \in \Gamma(R)$. For any $x, y \in X$,

$$\begin{aligned} \overline{(A \oplus a)}(x) \oplus \overline{R}(x, y) &= \overline{A}(x) \oplus a \oplus \overline{R}(x, y) = \overline{A}(x) \oplus \overline{R}(x, y) \oplus a \\ &\geq \overline{A}(y) \oplus a = \overline{(A \oplus a)}(y). \end{aligned}$$

Then

$$\overline{(A \oplus a)}(y) \ominus \overline{(A \oplus a)}(x) \leq \overline{R}(x, y).$$

We have $A \oplus a \in \Gamma(R)$.

$$(\overline{A}(y) \ominus \overline{A}(x)) \oplus \overline{(A \ominus a)}(x) = (\overline{A}(y) \ominus \overline{A}(x)) \oplus (\overline{A}(x) \ominus a)$$
$$\geq \overline{A}(y) \ominus a = \overline{(A \ominus a)}(y).$$

So

$$\overline{(A \ominus a)}(y) \ominus \overline{(A \ominus a)}(x) \leq (\overline{A}(y) \ominus \overline{A}(x)) \leq \overline{R}(x, y).$$

Therefore, $A \ominus a \in \Gamma(R)$.

Theorem 3 *Let $\mathcal{T}$ be an $\overline{L}$-topology on X. For any $x, y \in X$, set*

$$\overline{\Omega(\mathcal{T})}(x, y) = \bigvee_{\mu \in \mathcal{T}} \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right).$$

then $\Omega(\mathcal{T})$ is an $\overline{L}$-preorder on X.

Proof For any $x, y, z \in X$,

$$\overline{\Omega(\mathcal{T})}(x, x) = \bigvee_{\mu \in \mathcal{T}} \left(\overline{\mu}(x) \ominus \overline{\mu}(x)\right) = 0.$$

$$\begin{aligned}\overline{\Omega(\mathcal{T})}(x, y) \oplus \overline{\Omega(\mathcal{T})}(y, z) &= \bigvee_{\mu \in \mathcal{T}} \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right) \oplus \bigvee_{\mu \in \mathcal{T}} \left(\overline{\mu}(z) \ominus \overline{\mu}(y)\right) \\ &\geq \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right) \oplus \left(\overline{\mu}(z) \ominus \overline{\mu}(y)\right) \\ &\geq \overline{\mu}(z) \ominus \overline{\mu}(x) \\ &= \overline{\Omega(\mathcal{T})}(x, z).\end{aligned}$$

Theorem 4 *Let R be an $\overline{L}$-preorder on X. Then $R = \Omega \circ \Gamma(R)$.*

Proof For any $x, y \in X$,

$$\overline{\Omega \circ \Gamma(R)}(x, y) = \bigvee_{\mu \in \Gamma(R)} \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right) \leq \overline{R}(x, y)$$

Conversely, set an $\overline{L}$-subset $[x]_R$ of X as

$$\overline{[x]_R}(z) = \overline{R}(x, z), \quad \text{for any } z \in X.$$

Then $[x]_R \in \Gamma(R)$. Therefore,

$$\begin{aligned}\overline{\Omega \circ \Gamma(R)}(x, y) &= \bigvee_{\mu \in \Gamma(R)} \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right) \\ &\geq \overline{[x]_R}(y) \ominus \overline{[x]_R}(x) = \overline{R}(x, y) \ominus \overline{R}(x, x) = \overline{R}(x, y) \ominus 0 \\ &= \overline{R}(x, y).\end{aligned}$$

Theorem 5 *Let $\mathcal{T}$ be an $\overline{L}$-topology on X. Then $\mathcal{T}$ is Alexandroff and strong if and only if $\mathcal{T} = \Gamma \circ \Omega(\mathcal{T})$.*

Proof (ii) The sufficiency is obvious. As to the necessity, let $\mathcal{T}$ be an Alexandroff strong $\overline{L}$-topology on X.

If $A \in \mathcal{T}$, then for any $x, y \in X$,

$$\overline{A}(y) \ominus \overline{A}(x) \leq \bigvee_{\mu\in\mathcal{T}} \left(\overline{\mu}(y) \ominus \overline{\mu}(x)\right) = \overline{\Omega(\mathcal{T})}(x, y).$$

Therefore, $A \in \Gamma \circ \Omega(\mathcal{T})$.

Conversely, suppose $A \in \Gamma\circ\Omega(\mathcal{T})$. For any fixed $x \in X$, define an $\overline{L}$-subset m_x of X as follow:

$$\overline{m_x}(z) = \overline{A}(x) \oplus \overline{\Omega(\mathcal{T})}(x, z), \quad \text{for any } z \in X.$$

Then $\overline{m_x}(x) = \overline{A}(x)$ and $\overline{m_x}(z) \geq \overline{A}(z)$ for any $z \in X$. Therefore, $A = \bigcup_{x\in X} m_x$.

For any $\mu \in \mathcal{T}$, define an $\overline{L}$-subset g_μ of X as follow:

$$\overline{g_\mu}(z) = \overline{\mu}(z) \ominus \overline{\mu}(x), \quad \text{for any } z \in X.$$

Then $g_\mu \in \mathcal{T}$ by (O5). Since $\mathcal{T}$ is Alexandroff, we have $\bigcap_{\mu\in\mathcal{T}} g_\mu \in \mathcal{T}$.

Notice that

$$\begin{aligned}\overline{m_x}(z) &= \overline{A}(x) \oplus \overline{\Omega(\mathcal{T})}(x, z) = \overline{A}(x) \oplus \bigvee_{\mu\in\mathcal{T}} \left(\overline{\mu}(z) \ominus \overline{\mu}(x)\right) \\ &= \overline{A}(x) \oplus \bigvee_{\mu\in\mathcal{T}} \overline{g_\mu}(z) = \overline{A}(x) \oplus \overline{\bigcap_{\mu\in\mathcal{T}} g_\mu}(z),\end{aligned}$$

Therefore, $m_x \in \mathcal{T}$ by (O4) and $A = \bigcup_{x\in X} m_x \in \mathcal{T}$ by (O3).

By virtue of Theorems 4 and 5, we have the following conclusion.

Theorem 6 *There is a one-to-one correspondence between the set of all Alexandroff strong $\overline{L}$-topologies on X and that of all $\overline{L}$-preorders on X.*

4 Conclusion

In this paper, the concepts of $\overline{L}$-topological spaces and $\overline{L}$-preorders are introduced where L is a complete co-residuated lattice, by virtue of the degree of nonmembership in intuitionistic fuzzy set theory. A given $\overline{L}$-preorder on a set X can induce an

$\overline{L}$-topology on X, and vice versa. It is shown that there is a one-to-one correspondence between the set of all Alexandroff strong $\overline{L}$-topologies on X and that of all $\overline{L}$-preorders on X.

Acknowledgments This work is supported by the Natural Science Foundation of China (11401547) and the Fundamental Research Funds for the Central Universities (201413015).

References

1. Bělohlávek, R.: Fuzzy Relational Systems: Foundations and Principles. Kluwer Academic, Plenum Publishers, New York (2002)
2. Chang, C.L.: Fuzzy topological spaces. J. Math. Anal. Appl. **24**, 37–42 (1968)
3. Chen, P., Zhang, D.: Alexandroff *L*-co-topological spaces. Fuzzy Sets Syst. **161**, 2505–2514 (2010)
4. Chen, P., Lai, H., Zhang, D.: Coreflective hull of finite strong *L*-topological spaces. Fuzzy Sets Syst. **182**, 79–92 (2011)
5. Fang, J., Chen, P.: One-to-one correspondence between fuzzifying topologies and fuzzy preorders. Fuzzy Sets Syst. **158**, 1814–1822 (2007)
6. Flagg, R.C.: Quantales and continuity spaces. Algebra Univers. **37**, 257–276 (1997)
7. Gierz, G., Hofmann, K.H., Keimel, K., Lawson, J.D., Mislove, M.W., Scott, D.S.: Continuous Lattices and Domains. Cambridge University Press, Cambridge (2003)
8. Höhle, U.: Fuzzy real numbers as Dedekind cuts with respect to a multiple-valued logic. Fuzzy Sets Syst. **24**, 263–278 (1987)
9. Höhle, U.: Commutative, residuated l-monoids. In: Höhle, U., Klement, E.P. (eds.) Non-classical Logics and Their Applications to Fuzzy Subsets: A Handbook on the Mathematical Foundations of Fuzzy Set Theory, pp. 53–105. Kluwer Academic Publishers, Dordrecht (1995)
10. Lai, H., Zhang, D.: Fuzzy preorder and fuzzy topology. Fuzzy Sets Syst. **157**, 1865–1885 (2006)
11. Liu, Y., Luo, M.: Fuzzy Topology. World Scientific Publishing, Singapore (1997)
12. Rosenthal, K.I.: Quantales and Their Applications. Longman Scientific and Technical (1990)
13. Wang, G.J.: The theory of *L*-fuzzy topological spaces. Shanxi Normal University Press, Shanxi (1988) (in Chinese)
14. Yao, W., Shi, F.: A note on specialization L-preorder of L-topological spaces, L-fuzzifying topological spaces, and L-fuzzy topological spaces. Fuzzy Sets Syst. **159**, 2586–2595 (2008)
15. Ying, M.: A new approach to fuzzy topology I. Fuzzy Sets Syst. **39**, 303–321 (1991)
16. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338C–353 (1965)
17. Zhang, D.: An enriched category approach to many valued topology. Fuzzy Sets Syst. **158**, 349–366 (2007)

A Classification Method of Fuzzy Sets Based on Rough Fuzzy Number

X. Liu, K. Liang, Z. Liu and S. Wang

Abstract In view of the fuzzy classification problem, based on the theory of rough sets and fuzzy sets, we established a way of approaching fuzzy number—rough fuzzy number. Then a classification judgment function is established according to the classification of the two principles, it is used to evaluate classification effectively. Under the guidance of this idea, we got a fuzzy classification method. The paper also gives a application example, shows the whole process of taxonomy, to facilitate understanding and application.

Keywords Fuzzy classification · Rough fuzzy number · Fuzzy classification judgment function

1 Introduction

The classification of fuzzy sets is a complex uncertainty problem. Due to the fuzzy set is describing uncertain problems, so the classification problems has the characteristics of the unique solution. This kind of problem is complicated and is one of hot research issues.

Fuzzy classification should starting from two aspects, one is the study of the fuzzy sets, the other is a study of classification criteria. In this paper, using the theory of rough set, a fuzzy sets is established in general form—rough fuzzy number. It is a kind of approximation of fuzzy sets, but also has the characteristics of convenient operation, this opens the door for the application of the fuzzy set

The key of classification problem is to establish classification criteria. Combining with the fuzzy mean value method, a classification evaluation function is established, and completed the classification of fuzzy sets.

X. Liu · K. Liang · Z. Liu (✉) · S. Wang
College of Information Engineering, Dalian University, Dalian 116622, China
e-mail: liuzixin@dlu.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_4

2 Rough Fuzzy Number

In order to simplify the problem, assuming a fuzzy set $\tilde{A}$ is a regular convex fuzzy sets on [0,1]. Obviously, the transformation from the real number domain to the unit interval is easier, this kind of $\tilde{A} \in \tilde{F}\{[0,1]\}$ won't lose the general.

Let $[0,1]/R = \{[0,l_1],(l_1,l_2],\ldots,(l_{n-2},l_{n-1}],(l_{n-1},l_n]\} = \{X_1,X_2,\ldots,X_n\}$

Definition 1 Assume $([0,1],R)$ is a approximate space, for $\tilde{A} \in \tilde{F}\{[0,1]\}$

Let $\underline{\alpha}_i = \underline{R}(\tilde{A})(x) = \min\{\tilde{A}(y)\,|\,y \in X_i = [x]_R\}$

$$\overline{\alpha_i} = \overline{R}(\tilde{A})(x) = \max\{\tilde{A}(y)\,|\,y \in X_i = [x]_R\}$$

$\underline{R}(\tilde{A})$ is called lower approximation of fuzzy set $\tilde{A}$ according to $([0,1],R)$, similarly, $\bar{R}(\tilde{A})$ is called upper approximation of fuzzy set $\widetilde{A}$ according to $([0,1],R)$.

It is not difficult to verify, $\underline{R}(\tilde{A}) \subseteq \tilde{A} \subseteq \bar{R}(\tilde{A})$, or $\underline{R}(\tilde{A})(x) \le \mu_{\tilde{A}}(x) \le \bar{R}(\tilde{A})(x)$, $\mu_{\tilde{A}}(x)$ is the membership function of fuzzy set $\tilde{A}$.

Definition 2 Let $F^R_{\tilde{A}}$ is a rough fuzzy number, $x \in X_i, i = 1,2,\ldots,n$, $\underline{\alpha}_i \le F^R_{\tilde{A}} \le \overline{\alpha}_i$ and $F^R_{\tilde{A}}(x)$ is a constant. $F^R_{\tilde{A}}$ is a Step function of $[0,1]$, it is a constant when $x \in X_i$, $i = 1,2,\ldots,n$.

Thus $F^R_{\tilde{A}}$ is an approaching form of fuzzy set $\tilde{A}$, and $F^R_{\tilde{A}}$ can be represented as a vector form in determined approximation space $([0,1],R)$ [2].

$$F^R_{\tilde{A}}(x) = \cup^n_{i=1} f_i X_i,\; f_i = F^R_{\tilde{A}}(x),\; x \in X_i$$
$$F^R_{\tilde{A}}(x) = \{f_1,f_2,\ldots,f_n\}$$

The support of $F^R_{\tilde{A}}(x)$ is same with $\overline{R}(\tilde{A})$, the kernel of $F^R_{\tilde{A}}(x)$ is same with $\underline{R}(\tilde{A})$, $F^R_{\tilde{A}}(x)$ is a constant between $\underline{\alpha}$ and $\bar{\alpha}$, As the membership degree of $\widetilde{A}(x)$ in X_i, for example, let $\frac{\underline{\alpha}+\bar{\alpha}}{2} = F^R_{\tilde{A}}(x)$, $i = 1,2,\ldots,n$.

Theorem 1 *In a given approximation space,* $F^R_{\tilde{A}}$ *has the following properties:*

(1) $\forall \tilde{A},\tilde{B} \in \tilde{F}\{[0,1]\}$, *if* $\tilde{A} \subseteq \tilde{B}$, *then* $F^R_{\tilde{A}} \subseteq F^R_{\tilde{B}}$;
(2) $\forall \tilde{A},\tilde{B} \in \tilde{F}\{[0,1]\}$, *if* $\tilde{A} = \tilde{B}$, *then* $F^R_{\tilde{A}} = F^R_{\tilde{B}}$;
But instead is not necessarily true.
When $F^R_{\tilde{A}}(x) = \frac{\underline{\alpha}+\bar{\alpha}}{2}$, *we have*:
(3) $F^R_{\tilde{A}\cup\tilde{B}} \supseteq F^R_{\tilde{A}} \cup F^R_{\tilde{B}}$

Proof For a given approximation space $([0,1],R)$, for $\forall \tilde{A},\tilde{B} \in \tilde{F}\{[0,1]\}$, there is always $\tilde{A} \cup \tilde{B} \supseteq \tilde{A}, \tilde{A} \cup \tilde{B} \supseteq \tilde{B}, F^R_{\tilde{A}\cup\tilde{B}} \supseteq F^R_{\tilde{A}}, F^R_{\tilde{A}\cup\tilde{B}} \supseteq F^R_{\tilde{B}}$, so there are $F^R_{\tilde{A}\cup\tilde{B}} \supseteq F^R_{\tilde{A}} \cup F^R_{\tilde{B}}$.

(4) $F^R_{\tilde{A}\cap\tilde{B}} \subseteq F^R_{\tilde{A}} \cap F^R_{\tilde{B}}$

Proof For a given approximation space $([0,1],R)$ and $\forall\tilde{A},\tilde{B}\in\tilde{F}\{[0,1]\}$, there is always $\tilde{A}\cap\tilde{B}\subseteq\tilde{B}$, so when $F^R_{\tilde{A}\cap\tilde{B}}\subseteq F^R_{\tilde{A}}$, $F^R_{\tilde{A}\cap\tilde{B}}\subseteq F^R_{\tilde{B}}$, therefor $F^R_{\tilde{A}\cap\tilde{B}}\subseteq F^R_{\tilde{A}}\cap F^R_{\tilde{B}}$[3].

3 Classification Method of Fuzzy Sets and Classification Judgment Function

Let $\widetilde{Y}_1, \widetilde{Y}_2, \ldots \widetilde{Y}_n$ be fuzzy sets, $\widetilde{Y_{n_i}} \in \widetilde{F}\{[0,1]\}, i=1,2,\ldots,n, ([0,1],R)$, is a given approximation space, from previous chapter we get rough fuzzy numbers: $F^R_{\underset{Y_1}{\sim}}, F^R_{\underset{Y_2}{\sim}}, \ldots, F^R_{\underset{Y_n}{\sim}}$, and $F^R_{\underset{Y_i}{\sim}} = \{f_{i_1}, f_{i_2}, \ldots, f_{i_m}\}, 0\le f_{i_k}\le 1, k=1,2,\ldots,m.$

Each fuzzy set is written as approaching form of m dimensional vector, by the ideas of the geometric proximity, we classificate n fuzzy sets:

Let Classification number c is determined, and assume the initial classification of n fuzzy sets is k, remark as $F^R_{X_1}\ldots F^R_{X_k}$, simply remark as $X_1, X_2, \ldots, X_k, X_i = \{x_{i_1}, x_{i_2}, \ldots, x_{i_m}\}, i=1,2,\ldots,k.$

When $k>c$, then let k fuzzy sets be c class is called as clustering process;

When $k<c$, then let k fuzzy sets be c class is called as Classification process.

When $k\ge c$, clustering process is:

Step 1 Select a subclass center $V_i (i=1,2,\ldots,k)$ for every subclass;

Assume for subclass s: $\{X_{i_1}, X_{i_2}, \ldots, X_{i_s}\}$, we have $V_{i_t} = \frac{\sum_{j=1}^{s} x_{i_j t} x_{i_j t}}{\sum_{j=1}^{s} x_{i_j t}}, t = 1,2,\ldots,m.$

Step 2 Calculate each subclass of each element and the subclass center distance (according to the Euclidean distance), and sum, remark as $\varphi_i, i=1,2,\ldots,k.$

Step 3 Calculate the distance between each subclass center, remark as ψ_{ij}, denote the distance between i and j subclass.

Step 4 Calculate $M_\varphi = \max\{\varphi_i\}, \min\{\varphi_i\} = m_\varphi, M_\psi = \max\{\psi_{ij}\}, \min\{\psi_{ij}\} = m_\psi.$

Step 5 Introduct classification judgment function $LL(\varphi,\psi)(1<c<n)$:

$$LL(\varphi,\psi) = \sum_{i=1}^{k} \frac{M_\varphi - \varphi_i}{M_\varphi - m_\varphi} + \sum_{i<j} \frac{\psi_{ij} - m_\psi}{M_\psi - m_\psi}.$$

Calculate $LL(\varphi,\psi)$, the maximum as the classification mode of $c=k$.

When $k>c$, select $m_\psi = \psi_{st}$, let X_s and X_t as aggregate object, from step 1 to step 5 again, get new $LL(\varphi,\psi)$, untill $c=k$ stop.

Similarly, when $k < c$, select $M_\varphi = \varphi_s$, let X_s as classification object, Divided subclass X_s into two new subclasses $X_s^{'}$ and $X_s^{''}$, then calculate new $LL(\varphi, \psi)$, untill $c = k$ stop.

When $c = 2$, in $\sum\limits_{i<j} \frac{\psi_{ij} - m_\psi}{M_\psi - m_\psi}$, ψ_{ij} only have ψ_{12}, then $\sum\limits_{i=j} \frac{\psi_{ij} - m_\psi}{M_\psi - m_\psi}$ is ψ_{12}.

When $c = 1$, Classification of the object as a whole, that there is no classification, at this moment, $LL(\varphi, \psi)$ as the distance sum of classification object and its center.

When $c = n$, every classification object as a subclass, denote that classification arrive its limits, at this moment, preceding paragraph of $LL(\varphi, \psi)$ is zero, $LL(\varphi, \psi)$ is equal to the distance sum between n objects.

$LL(\varphi, \psi)$ Consists of two item, The first item is geometric proximity between the elements for the subclass, describe classification from local, the index as small as possible. The second item is geometric proximity among the subclass, describe classification from global, the index is larger, the better. So $LL(\varphi, \psi)$ 从compre-hensively evaluate classification from the local and global. we can Introduct preference factor λ according to the actual situation.Then $LL(\varphi, \psi) = \lambda \sum\limits_{i=1}^{k} \frac{M_\varphi - \varphi_i}{M_\varphi - m_\varphi} + (1 - \lambda) \sum\limits_{i<j} \frac{\psi_{ij} - m_\psi}{M_\psi - m_\psi}, 0 < \lambda < 1$.

Deciside the balance of local and global relations by $LL(\varphi, \psi)$, get satisfactory classification model Finally [4].

3 An example of fuzzy classification analysis

Let fuzzy classification sample set $\widetilde{X} = \{\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3, \widetilde{X}_4\}, \widetilde{X}_i (i = 1, 2, 3, 4)$, is triangular fuzzy number. And $\widetilde{X}_1 = (0, 0.4, 0.9); \widetilde{X}_i (i = 1, 2, 3, 4); \widetilde{X}_3 = (0.4, 0.8, 0.9); \widetilde{X}_4 = (0.6, 0.85, 0.95)$.

Divided $\{\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3, \widetilde{X}_4\}$ into $c = 2$.

$$\widetilde{X}_1 = \begin{cases} 2.5x & 0 \le x < 0.4 \\ 1 - 2(x - 0.4) & 0.4 \le x < 0.9 \\ 0 & 0.9 \le x \le 1 \end{cases} \qquad \widetilde{X}_2 = \begin{cases} 2x & 0 \le x < 0.5 \\ 1 - \frac{10}{3}(x - 0.5) & 0.5 \le x < 0.8 \\ 0 & 0.8 \le x \le 1 \end{cases}$$

$$\widetilde{X}_3 = \begin{cases} 0 & 0 \le x < 0.4 \\ 2.5(x - 0.4) & 0.4 \le x < 0.8 \\ 1 - 10(x - 0.8) & 0.8 \le x < 0.9 \\ 0 & 0.9 \le x \le 1 \end{cases} \qquad \widetilde{X}_4 = \begin{cases} 0 & 0 \le x < 0.6 \\ 4(x - 0.6) & 0.6 \le x < 0.85 \\ 1 - 10(x - 0.85) & 0.85 \le x < 0.95 \\ 0 & 0.95 \le x \le 1 \end{cases}$$

Divided $[0, 1]$ into seven interval:

$$[0, 0.35), [0.35, 0.45), [0.45, 0.55), [0.55, 0.65), [0.65, 0.85), [0.85, 0.9), [0.9, 1]$$

We get rough fuzzy sets:

$$\underline{\widetilde{X_1}} = \{0, 0.88, 0.9, 0.7, 0.5, 0.1, 0\}, \overline{\widetilde{X}}_1 = \{0.88, 0.9, 0.7, 0.5, 0.1, 0, 0\}$$
$$\underline{\widetilde{X_2}} = \{0, 0.7, 0.9, 0.83, 0.5, 0, 0\}, \overline{\widetilde{X}}_2 = \{0.7, 0.9, 0.83, 0.5, 0, 0, 0\}$$
$$\underline{\widetilde{X_3}} = \{0, 0, 0.13, 0.38, 0.63, 0.5, 0\}, \overline{\widetilde{X}}_3 = \{0, 0.13, 0.38, 0.63, 0.5, 0, 0\}$$
$$\underline{\widetilde{X_4}} = \{0, 0, 0, 0, 0.2, 1, 0.5\}, \overline{\widetilde{X}}_4 = \{0, 0, 0, 0.2, 1, 0.5, 0\}$$

Calculation of rough fuzzy number is:

$$F^R_{\underset{\sim}{X_1}} = \{0.44, 0.89, 0.8, 0.6, 0.3, 0.05, 0\}, F^R_{\underset{\sim}{X_2}} = \{0.35, 0.8, 0.87, 0.67, 0.25, 0, 0\},$$
$$F^R_{\underset{\sim}{X_3}} = \{0, 0.07, 0.26, 0.51, 0.57, 0.25, 0\}, F^R_{\underset{\sim}{X_4}} = \{0, 0, 0, 0.1, 0.6, 0.75, 0.25\}$$

At present $c = 4$, each $\widetilde{X}_i$ as a kind, calculate the distance between them, $d_{ij} = \left\| F^R_{\underset{\sim}{X_i}} F^R_{\underset{\sim}{X_j}} \right\|$ denote the distance between $\widetilde{X}_i$ and $\widetilde{X}_j$.

$$d_{12} = 0.1688, d_{13} = 1.1129, d_{14} = 1.4843, d_{23} = 1.0749$$
$$d_{24} = 1.4889, d_{34} = 0.6550$$

From the above results we find d_{12} is the minimum, then we combine $\widetilde{X}_1$ and $\widetilde{X}_2$ as a kind, get three kind: $\{\widetilde{X}_1, \widetilde{X}_2\}, \widetilde{X}_3, \widetilde{X}_4$. Then according to the classification step, we have:

Step 1 calculate the first subclass center: $V_{1k} = \frac{\sum_{i=1}^{2} X_{ik} \cdot X_{ik}}{\sum_{i=1}^{2} X_{ik}}$

$$F^R_{\underset{\sim}{X_1}} = \{0.44, 0.89, 0.8, 0.6, 0.3, 0.05, 0\}, F^R_{\underset{\sim}{X_2}} = \{0.35, 0.8, 0.87, 0.67, 0.25, 0, 0\}$$
$$V_{11} = 0.4, V_{12} = 0.847, V_{13} = 0.836, V_{14} = 0.637, V_{15} = 0.277$$
$$V_{16} = 0.05, V_{17} = 0$$

The new classification of $\widetilde{X}_1$ and $\widetilde{X}_2$ is $\widetilde{X}_1'$, its center is $V^{\widetilde{X}_1'}$, Shorthand for $V^{(1)}$, $V^{(1)} = \{0.4, 0.847, 0.836, 0.637, 0.277, 0.05, 0\}$.

Step 2 calculate φ_1

The distance between $\widetilde{X}_1$ and $\widetilde{X}_1'$ is denoted by d_1', The distance between $\widetilde{X}_2$ and $\widetilde{X}_1'$ is denoted by d_2': $d_1' = 0.0687$, $d_2' = 0.0769$.

Step 3 calculate $\psi_{13}, \psi_{14}, \psi_{34}$

$$\psi_{13} = 1.1129, \psi_{14} = 1.5883, \psi_{34} = d_{34} = 0.6550$$

From the above results we find the distance between $\widetilde{X}_3$ and $\widetilde{X}_4$ is the minimum, combine $\widetilde{X}_3$ and $\widetilde{X}_4$, then $\widetilde{X}_3' = \{\widetilde{X}_3, \widetilde{X}_4\}$, $\widetilde{X}_1' = \{\widetilde{X}_1\}, \widetilde{X}_2' = \{\widetilde{X}_2\}$, Then according to the classification step, we have:

Step 1 calculate the subclass center of $\widetilde{X}_3'$, denoted by $V_3 = \{v_{31}, v_{32}, \ldots, v_{37}\}$

$$v_{3k} = \frac{\sum_{i=3}^{4} x_{ik} \cdot x_{ik}}{\sum_{i=3}^{4} x_{ik}}, k = 1, 2, \ldots, 7$$

$$v_{31} = 0, v_{32} = 0.07, v_{33} = 0.26, v_{34} = 0.4428, v_{35} = 0.5854, v_{36} = 0.625, v_{37} = 0.25$$

Step 2 calculate $\varphi_3(\varphi_1 = 0, \varphi_2 = 0)$
The distance between $\widetilde{X}_3$ and $\widetilde{X}_3'$ is denoted by d_3', The distance between $\widetilde{X}_4$ and $\widetilde{X}_3'$ is denoted by d_4'.

$$d_3' = 0.4559, d_4' = 0.4537$$

Step 3 calculate $\psi_{12}', \psi_{13}', \psi_{23}'$ (among them, $\psi_{12}' = d_{12} = 0.1688$)
$\psi_{13}' = 1.2872, \psi_{23}' = 1.2825$ combine $\widetilde{X}_1'$ and $\widetilde{X}_2'$, then we get two classification $\widetilde{X}_1'' = \{\widetilde{X}_1, \widetilde{X}_2\}$ and $\widetilde{X}_2'' = \{\widetilde{X}_3, \widetilde{X}_4\}$.

Step 1 calculate the center of two subclass:

$$v_1'' = \{0.4, 0.847, 0.836, 0.637, 0.277, 0.05, 0\}$$
$$v_1'' = \{0, 0.07, 0.26, 0.4428, 0.5854, 0.625, 0.25\}$$

Step 2 calculate φ_1 and φ_2, $d_i^{(j)}$ denote the distance between $\widetilde{X}_i$ and v_j. (that is, $d_1^{(1)}$ denote the distance between $\widetilde{X}_1$ and $v_1, d_2^{(1)}, d_3^{(2)}, d_4^{(2)}$, is similarly)

$$d_1^{(1)} = 0.0687, d_2^{(1)} = 0.0709, d_3^{(2)} = 0.4559, d_4^{(2)} = 0.4537$$
$$\varphi_1 = 0.0687 + 0.0709 = 0.1396, \varphi_2 = 0.4559 + 0.4537 = 0.9096$$

Step 3 calculate $\psi_{12}, \psi_{12} = 1.2738$

Step 4 calculate $M_\varphi, m_\varphi, M_\psi, m_\psi$

$$M_\varphi = 0.9096, m_\varphi = 0.1396, M_\psi = 1.2738, m_\psi = 1.2738$$

Step 5 calculate $LL(\varphi, \psi)$

$$LL(\varphi, \psi) = 1 + 1.2738 = 2.2738$$

To verify the other classification: (please verify $\widetilde{X}_1' = \{\widetilde{X}_1\}, \widetilde{X}_2' = \{\widetilde{X}_2, \widetilde{X}_3, \widetilde{X}_4\}$)

It is easy to verify, when $c = 2$, other classification combination Less than 2.2738, By the principle of the optimal, the classification results are $\{\widetilde{X}_1, \widetilde{X}_2\}, \{\widetilde{X}_3, \widetilde{X}_4\}$.

4 Conclusion

The method proposed in this paper approximating a fuzzy set by the theory of rough fuzzy set. This method is a complement of fuzzy classification. It is a kind of exploration and attemption according to fuzzy classification problem and provides a new thought.

Fuzzy approximation covering the partition of approximation space etc., its affection to fuzzy approximation is the further topics should be discussed; About classification under the control of local and global balance also has a lot of content to talk about [6]. This article is only a kind of balance and using the method, there will be more better way in practice, this is what we are looking forward to.

Acknowledgement Thanks to the support by National Natural Science Foundation of China (No.61170255)

References

1. Kapp, A.V., Tibshirani, R.: Are clusters found in one dataset present in another dataset. Biostatistics **8**(1), 9–31 (2007)
2. Yu, H., Luo, H.: A novel possibilistic fuzzy leader clustering algorithm. RSFDGrC 2009
3. Boixader, D., Jaces, J., Recasens, J.: Upper and lower approximations of fuzzy sets. Int. J. Gen. Syst. **29**, 555–568 (2000)
4. Liu, X., Wu, W., Hu, J.: A method of fuzzy multiple attribute decision making based on rough sets. IcIc Int. J. **4**(8), 2005–2010 (2008)
5. Shao-Hui, L., Fei, H., Zi-Yan, J., Zhong-Zhi, S.: A rough set-based hierarchical clustering algorithm. J. Comput. Res. Dev. **41**(4), 552–557 (2004)
6. Liu, D., Li, H., Zhou, X.: Two decades research on decision-theoretic rough sets 11. In: Proceedings of 9th IEEE International Conferences on Cognitive Informatics (2010)
7. Yu, H., Chu, S., Yang, D.: A semiautonomous clustering algorithm based on decision-theoretic tough set theory 11. In: The 9th IEEE international Conference on Cognitive Informatics (2010)

A Prediction Model for Hot Metal Desulfurization Rate Based on Fuzzy Structured Element

Jun-hong Ji, Ru-quan Liang and Ji-cheng He

Abstract The desulfurization quantity is a fuzzy number in the hot metal desulfurization process. The fuzzy multiple linear regression model was established using the fuzzy structure element and multiple linear regression theory, in which the fuzzy desulphurization quantity was related to the original condition of hot metal, the composition of desulfurizer, and injection operational parameters. The fuzzy desulfurization rate was used to measure the fuzzy desulfurization quantity, and 3000 group desulfurization data of hot metal from Anshan Iron and Steel were used for the empirical study. The results show that the present model can provide decision-makers with richer information, which can express the uncertain relationship between parameters of hot metal desulfurization and the fuzzy desulfurization quantity with a lower average absolute error.

Keywords Hot metal desulfurization rate · Forecasting model · Fuzzy multiple linear regression model · Structural element

1 Introduction

The hot metal desulphurization is an essential process for improving the product quality and the production efficiency. If the final sulfur content and target sulfur content have big difference after desulphurization, which not only increases the burden of converter and other refining equipment, but also causes the sulfur content to exceed the standard, which can not meet the requirements of steel types. So the study

J. Ji · R. Liang (✉) · J. He
Key Laboratory of Electromagnetic Processing of Materials (Ministry of Education), Northeastern University, Shenyang 110819, China
e-mail: liang@epm.neu.edu.cn

J. Ji
School of Safety Science and Engineering, Liaoning Technical University, Huludao 125105, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_5

of forecasting model on the effect of the hot metal desulphurization can provide more accurate desulfurization information, thus providing guidance for steel production.

So far, scholars have conducted a lot of researches about the hot metal desulphurization process. In Refs. [1–3], in which radial basis function neural network (RBF-RNN), back-propagation neural network (BP-RNN) etc. were applied to improve the forecasting model to predict the desulfurization quantity of injection desulfurization, respectively. In Ref. [4], basis on BP neural network forecasting model, the neural network forecasting model was established for desulfurization according to the gray correlation degree in the gray theory. In Ref. [5], the parameters controlling model was established during the injection process according to the experience data of engineering practice, which achieved automatic calculation and precise control of process parameters. In Refs. [6, 7], the comparison expression between process parameters and the desulfurization rate was established by applying the statistical principles for regression, through a large number of statistics about the practical data of enterprise production, and the corresponding control software about the desulfurizer consumption was developed. In Refs. [8, 9], the reaction essence was studied and the desulfurization rate was discussed from the desulfurization mechanism. The above researches can promote the cognition to the factors affecting the desulfurization quantity effectively and improving the accuracy effect of forecasting desulphurization.

The final sulfur content of hot metal is a important index for enterprises to measure the desulfurization effect. However, in production practice, final sulfur content is different from the target sulfur content which set in theory, due to following reasons, different initial parameters of desulfurization, difficult-to-control operating parameters and some defects in the data acquisition and detection of final sulfur content. Furthermore, it is difficult to completely skim slag after injection, so resulfurization occurred in converter steelmaking process, the sulfur content in the resulfurization phenomenon is related to the result of hot metal desulfurization directly, and it is not calculated to the final sulfur content in the process of hot metal desulphurization. Therefore, the final sulfur content is an uncertain value. It is obvious that consider the final sulfur content as a fixed value, there is the strong possibility that the sulfur content in the subsequent working procedure exceeds the threshold and affects the steel quality. So taking the final sulfur content as an uncertain value is more practical.

Since the final sulfur content is an uncertain value, it is supposed as a fuzzy number. In practice, desulfurizing rate is usually used to measure the final sulfur content, and desulfurization rate is related to hot metals initial condition, the composition of desulfurizer, injection operation parameters, and the dosage of desulfurizer etc. So, present work regards the desulfurization rate as a fuzzy number, and uses fuzzy structure element theory to propose a fuzzy multiple linear regression model of hot metals desulphurization rate. Meanwhile, we use an example to prove the validity of the model, and solve the controlling problem of the desulfurizer quantity for production practice.

2 Preliminaries

2.1 *Fuzzy Structure Element and Its Expression*

Definition 1 [10] Set E as fuzzy set in the real number field R, and the membership function is $E(x), x \in R$. If $E(x)$ satisfies the following properties, $E(0) = 1$; in the interval $[-1, 0)$, the $E(x)$ is right continuously monotonic function, and in the interval $(0, 1]$, the $E(x)$ is left continuously monotonic function; when $x \in (-\infty, -1)$ or $x \in (1, +\infty)$, $E(x) = 0$, then E is called the fuzzy structure in the R.

Theorem 1 [12] Set E as any structure element, and the membership function is $E(x)$. Meanwhile let a function of $f(x)$ be monotonic and bounded in the interval $[-1, 1]$, then $f(E)$ is a fuzzy number in the R , and the membership function of $f(E)$ is $E(f^{-1}(x))$, in which $f^{-1}(x)$ is the symmetric alternating function of $f(x)$ about the variables x and y. If $f(x)$ is continuous and strictly monotonic function, then $f^{-1}(x)$ is the inverse function of $f(x)$.

Definition 2 Let R and L respectively be strictly decreasing continuous function in the $[0, +\infty]$ and $[0, 1]$ interval, and $L(0) = 1; x > 0, L(x) < 1; x < 1, L(x) > 0. L(1) = 0$ or $(\forall x \in R,\ L(x) > 0,\ L(+\infty) = 0)$, R and L are the same in kind. Then the fuzzy number $\tilde{A}$ is called LR- type fuzzy number, and if $a,\ \alpha,\ \beta \in R, \alpha > 0,\ \beta > 0$, then

$$\mu_{\tilde{A}}(x) = \begin{cases} L(\frac{a-x}{\alpha}),\ x \leqslant a, \\ R(\frac{x-a}{\beta}),\ x \geqslant a \end{cases} \tag{1}$$

The LR-type fuzzy number $\tilde{A}$ is remembered as $\tilde{A} = < a, \alpha, \beta >_{LR}$. When $L(x) = R(x) = 1 - x$, $\tilde{A} = < a, \alpha, \beta >_{LR}$ is the triangular fuzzy number, which is denoted as $\tilde{A} = < a, \alpha, \beta >$ in short.

Theorem 2 *Let E be the triangular structure element, and $\tilde{A}$ is the LR- type fuzzy number, and its membership function is the formula* (1)*, then there will exist a monotonic and bounded function f in the* $[-1, 1]$ *interval, and make $\tilde{A} = f(E)$, of which*

$$f(x) = \begin{cases} a - \alpha L^{-1}(1 + x),\ -1 \leqslant x \leqslant 0, \\ a + \beta R^{-1}(1 + x),\ 0 \leqslant x \leqslant 1. \end{cases} \tag{2}$$

2.2 *The Fuzzy Multiple Linear Regression Model*

Fuzzy multivariate linear regression model [10–12] can be used to study the relations between precise input and fuzzy output, and fuzzy input and fuzzy output. In practice, the model has good suitability and superiority for problems with multiple outputs. In forecasting hot metal desulfurization, the maximum possible desulfurization quantity and its possible interval can be obtained at the same time. And the information obtained will be more reliable, and its practical significance is more outstanding. According to the Ref. [12], Let inputs be precise numbers and outputs be fuzzy numbers among a set of observation data $(Z_i, \tilde{y}_i)$, in which $Z_i = (x_{1i}, x_{2i}, \ldots, x_{mi})$, $Z_i \in R^m$, $\tilde{y}_i \in \tilde{N}_C(R)$, $i = 1, 2, \ldots, n$, $n \geqslant 2$. Fuzzy linear regression model can be expressed as the following formula (3):

$$\tilde{y} = \tilde{a}_0 + \tilde{a}_1 x_1 + \tilde{a}_2 x_2 + \cdots + \tilde{a}_m x_m \tag{3}$$

In (3), $a_i \in \tilde{N}_C(R)$, $i = 0, 1, 2, \ldots, m$. And $\tilde{N}_C(R)$ are the whole of all boundary-closed fuzzy numbers in the real number set R. As can be shown in the formula (3), if $Z_i \geqslant 0$, for any $\tilde{y}_i$, the formula (4) can be obtained.

$$\overset{\wedge}{\tilde{y}}_i = (\hat{f}_0 + \hat{f}_1 x_{1i} + \hat{f}_2 x_{2i} + \cdots + \hat{f}_m x_{mi}) y|_{y=E} = \hat{g}_i(E) \tag{4}$$

In the formula (4), $g_i(y) = (f_0 + f_1 x_{1i} + f_2 x_{2i} + \cdots + f_m x_{mi})(y)$, $f_i \in B[-1, 1]$, $\tilde{a}_i = f_i(E)$. $B[-1, 1]$ is the standard monotonic and bounded function with all of the same order in the $[-1, 1]$ interval, E is the fuzzy structure element.

From the formula (4), estimated value $\overset{\wedge}{\tilde{y}}_i$ of $\tilde{y}_i$ can be expressed as the formula (5)

$$\overset{\wedge}{\tilde{y}}_i = (\hat{f}_0 + \hat{f}_1 x_{1i} + \hat{f}_2 x_{2i} + \cdots + \hat{f}_m x_{mi}) y|_{y=E} = \hat{g}_i(E) \tag{5}$$

In the formular (5), $\hat{g}_i(y) = (\hat{f}_0 + \hat{f}_1 x_{1i} + \hat{f}_2 x_{2i} + \cdots + \hat{f}_m x_{mi})(y)$, $\hat{f}_i(E)$ is the estimated parameter of $f_i(E)$ in (5), $\hat{f}_i \in B[-1, 1]$.

Thus on the basis of the fuzzy correlation coefficient (formula (5)), can be determined according to the formula (6)

$$\min \sum_{i=1}^{n} d_E^2(\tilde{y}_i, \overset{\wedge}{\tilde{y}}_i) = \min \sum_{i=1}^{n} d_E^2(g_i(E), \hat{g}_i(E)) \tag{6}$$

In the formula (6), the membership function of can be expressed as the formula (7).

$$u_{\overset{\wedge}{\tilde{y}}}(x_{1i}, x_{2i}, \ldots, x_{mi}; y) = E((\hat{f}_0 + \hat{f}_1 x_{1i} + \hat{f}_2 x_{2i} + \cdots + \hat{f}_m x_{mi})^{-1}(y)) \tag{7}$$

3 Fuzzy Multivariate Linear Regression Model of Desulfurization Rate

3.1 Problem Description and Model Establishment

The final sulfur content measured in the process of hot metal desulfurization has great influence on the subsequent steelmaking process. However, the final sulfur content is closely related to the operation parameters of injection and initial parameters of hot metal. In the desulfurization process, the detection is usually done after completing injection, to test the final sulfur content whether it meets the requirement or not. Because the desulfurizer injecting into the hot metal is a complicated gas-solid-liquid multiphase flow chemical reaction process, after completing the injection, there is still some desulfurizer reacting with the hot metal, after skim slag, it is inevitable to remain some residual slag in the hot metal, and the resulfurization can occur for those slag.

In order to get the interval in which the final sulfur content exists and the most possible values in this interval, the calculation model of the final sulfur content related to other parameters needs to be established. And the relevant simplifications conditions must be introduced into the theoretical analysis. So based on the actual production condition and operation experience of Anshan Iron and Steel Complex hot metal injection desulphurization, present paper establishes the calculation model of the injection desulphurization rate and injection parameter applying metallurgical reaction principle. Based on Refs. [7, 13], there exist a multielement linear relationship among desulphurization rate, temperature of hot metal, initial sulfur content of hot metal, and the dosage of desulfurizer and injection speed. Because the desulphurization rate is a fuzzy number, treating the relationship as fuzzy multivariate linear relationship is more grounded in the reality. So the relationship can be established as the formula (8).

$$\ln\tilde{\eta}_i^s = \tilde{a}_0 + \tilde{a}_1 \ln s_0 + \tilde{a}_2 \frac{1}{T} + \tilde{a}_3 \ln(I + 1) + \tilde{a}_4 \ln R_p^* + \tilde{a}_5 \ln G_p^* \tag{8}$$

In (8), s_0 is the initial sulfur content of hot metal; T is the temperature of hot metal at the beginning of desulphurization; I is the ratio of calcium and magnesium; R_p^* is the speed of injection; G_p^* is the quantity of injection; and $I = \frac{G_{CaO}^*}{G_{Mg}^*}$, $R_p^* = \frac{G_P^*}{t}$, $G_p^* = \frac{G_{CaO}+G_{Mg}}{Q} = G_{CaO}^* + G_{Mg}^*$, Kg/t.Fe; Q is the quantity of hot metal.

3.2 Solution of the Model

In order to express easily, make $\tilde{y}_i = \ln \eta_i^s$, $x_1 = \ln s_0$, $x_2 = \frac{1}{T_i}$, $x_3 = \ln(I + 1)$, $x_4 = \ln R_p^*$, $x_5 = \ln G_p^*$.

Make all of the fuzzy numbers involved in the model (8) LR- fuzzy numbers [12], namely: $\tilde{y}_i = < b_i, p_i, q_i >_{LR} = 2s_i - p_i L^{-1}(1 + E^-) + q_i R^{-1}(1 - E^+), i = 1, 2, \ldots, n$, $\tilde{a}_i = < a_i, \alpha_i, \beta_i >_{LR} = 2a_i - \alpha_i L^{-1}(1 + E^-) + \beta_i R^{-1}(1 - E^+), i = 1, 2, \ldots, m$. In which, $A = (a_0, a_1, a_2, \ldots a_m)^T$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \ldots \alpha_m)^T$, $\beta = (\beta_0, \beta_1, \beta_2, \ldots \beta_m)^T$, $X = (1, x_1, x_2, \ldots x_m)^T$, $B = (b_0, b_1, b_2, \ldots b_m)^T$, $P = (p_0, p_1, p_2, \ldots p_m)^T$, $Q = (q_0, q_1, q_2, \ldots q_m)^T$, $X_i = (1, x_{1i}, x_{2i}, \ldots x_{mi})^T$.
And then

$$\begin{aligned} \tilde{y} &= \tilde{a}_0 + \tilde{a}_1 x_1 + \tilde{a}_2 x_2 + \cdots + \tilde{a}_m x_m \\ &= < a_0, \alpha_0, \beta_0 >_{LR} + < a_1, \alpha_1, \beta_1 >_{LR} x_1 + \cdots + < a_m, \alpha_m, \beta_m >_{LR} x_m \\ &= 2A^T X - \alpha^T X \cdot L^{-1}(1 - E^-) + \beta^T X \cdot L^{-1}(1 - E^-) \end{aligned} \tag{9}$$

And then A, α, β, can be determined according to the Ref. [12], and they are shown in Eq. (10).

$$\begin{bmatrix} A \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} (X^T X)^{-1} X^T S \\ (X^T X)^{-1} X^T P \\ (X^T X)^{-1} X^T Q \end{bmatrix} \tag{10}$$

4 Calculation Examples

The desulfurization rate is related to the operating parameters of desulphurization and the initial parameters of hot metal in the process of hot metal desulfurization. The following is the relevant data collected about Calcium Magnesium compound injection desulphurization in Anshan Iron and Steel Complex and it includes 3000 groups, and 2966 groups among the data sets are effective, as shown in Table 1. Randomly select 2370 groups of data as experimental samples, and other 596 groups of

Table 1 Sample data

	x_1	x_2	x_3	x_4	x_5	$\tilde{y}_{2i}$	$\hat{\tilde{y}}_{2i}$
1	0.037	1345	3.295	14.293	45.1	<0.724,0.08, 0.06>	<0.724,0.063,0.045>
2	0.040	1343	3.314	13.304	46.6	<0.632,0.12, 0.04>	<0.633,0.09,0.032>
3	0.035	1324	3.673	10.185	47.2	<0.562,0.14, 0.08>	<0.563,0.12,0.065>
⋮				⋮			⋮
2370	0.058	1299	2.684	14.516	43.5	<0.754,0.07, 0.05>	<0.754,0.054,0.038>
2371	0.036	1343	3.670	7.669	46.7	<0.663,0.090.05>	<0.664,0.075,0.038>
2372	0.065	1297	2.610	12.519	47.3	<0.653,0.1, 0.04>	<0.654,0.08,0.026>
2373	0.072	1240	2.620	8.415	46.7	<0.768,0.06,0.02>	<0.768,0.054,0.015>
⋮				⋮			⋮
2966	0.048	1293	3.038	18.517	42	<0.784,0.08, 0.06>	<0.784,0.079,0.056>

Note The experimental samples are from 1 to 2370, and the test samples are from 2371 to 2966, the data are listed with three significant figures after the decimal point in the table data as samples for inspection. The desulphurization rate of hot metal on production site is calculated by the formula $\eta_s = \frac{s_0 - s}{s_0} \times 100\,\%$. The related variables x_1 is the natural logarithm of initial sulfur content of hot metal s_0 (%), and x_2 is the reciprocal of hot metals temperature $T\,(^oC)$ at the beginning of desulphurization, and x_3 is the natural logarithm of the sum of Ca-Ma ratio I and 1, and x_4 is the natural logarithm of the speed of injecting powder R_p (Kg/$t.Fe$ min), and x_5 is the natural logarithm of injection quantity of desulfurizer G_p (Kg/t.Fe).

4.1 Establishment and Calculation of Model

We record these operating parameters and desulfurization rate before and after the desulphurization of each car of hot metal ladle, and then summarize the data of desulfurization rate. If just replaced with a accurate value, much of the information will be lost, and the feasible method is to regard the desulfurization rate of each car of hot metal ladle as LR-fuzzy number $\tilde{y}_i =< b_i, p_i, q_i >_{LR}$, in which $i = 1, 2, \ldots, n$; b_i is the value of desulfurization rate of the order i hot metal ladle; p_i is the mean value that is less than b_i; q_i is the mean value that is more than b_i. The results are shown in Table 1.

It can be known from the second part in present paper, the fuzzy multiple regression model of hot metal desulphurization rate can be established like the formula (8). If recording the parameters of the ordermetal ladle as $Z_i = (x_{1i}, x_{2i}, \ldots, x_{mi})$, the desulphurization rate of the order metal ladle is $\tilde{y}_i =< b_i, p_i, q_i >_{LR}$, in which $i = 1, 2, \ldots, n$. And then by using, $A = (X^T X) X^T B, \alpha = (X^T X) X^T P, \beta = (X^T X) X^T Q$ the parameter estimation of fuzzy valued function will be obtained which are a group of LR-fuzzy numbers, namely $\hat{a}_0 =< 3.5956, -0.5647, 0.3663 >_{LR}$, $\hat{a}_1 =< 0.4538, -0.2687, 0.0512 >_{LR}$, $\hat{a}_2 =< 0.6642, -0.2569, 0.1255 >_{LR}$, $\hat{a}_3 =< -3.7336, 0.0066, -0.0785 >_{LR}$, $\hat{a}_4 =< -0.0405, 0.0213, -0.0011 >_{LR}$, $\hat{a}_5 =< 0.9405, 0.0221, -0.0731 >_{LR}$.

4.2 The Prediction and Test of Model

The model is trained by using experimental samples from 1 to 2370, and the results are shown in Table 1, based on the multiple linear regression model of desulphurization rate in hot metal desulfurization process. The average absolute error of desulfurization efficiency is 0.0349. It can be known that the model is reliable, and it can be applied in reality. According to the fuzzy model which has been trained, prediction is made for the test samples from 2371 to 2966, comparison is made with the actual observed values. The results are shown in Table 1. The average absolute error of desulfurization efficiency is 0.0274. Therefore, it is feasible by applying fuzzy structure element theory and regression analysis theory to predict the desulfurization efficiency of hot metals pretreatment with injection method.

It can be known from the test samples in Table 1 that the desulfurization rate obtained is a fuzzy value, which has the possible upper limit and lower limit, and then according to the formula of desulfurization rate, the variation range of desulfurization quantitycan be obtained. As an example of the No. 2966 sample, the predicted value of desulfurization rate is$< 0.784, 0.079, 0.056 >$. As can be known, the possibility that the predicted value of desulfurization rate in this furnace is 0.784 is the maximum, and the possibility that the value is not more than 0.705 or not less than 0.84 is 0; the possibility that the predicted value of desulfurization quantity is 0.010368 is the maximum, and its membership degree is

$$\mu = \begin{cases} 1 + \frac{x-0.010368}{0.002688}, 0.00768 \leqslant x \leqslant 0.010368, \\ 1 - \frac{x-0.010368}{0.003792}, 0.010368 < x \leqslant 0.1416. \end{cases}$$

5 Conclusions

The fuzzy multiple linear regression model was established by using the fuzzy structure element and multiple linear regression theory. The fuzzy desulfurization rate was used to measure the fuzzy desulfurization quantity, and 3000 group desulfurization data of hot metal were used for the empirical study. the present model can provide decision-makers with richer information and the fuzzy desulfurization quantity with a lower average absolute error.

(1) In this paper, based on fuzzy mathematics and multiple linear regression correlation theory, the desulphurization problem of hot metal was studied, the desulfurization rate obtained was not a concrete number, but an interval, and the corresponding possibility of each value occurrence was obtained, which was more grounded in reality.
(2) According to initial parameters and the target of hot metals desulfurization, the model can more accurately forecast the desulfurization rate and its fluctuation range, which is conducive to rapidly adjust the dosage of desulfurizer and get acclimatized to the change of complex working condition, improve production efficiency and improve the performance of steel.
(3) The desulphurization effect of hot metal has big uncertainty, and the uncertainty can be better expressed by fuzzy analysis theory. The solving processes is considered more comprehensively, and the content of results is more rich. But in the solving process of practical problems, the expression of triangular fuzzy number for fuzzy numbers should be paid attention to (some other expressions such as rectangular fuzzy number, trapezoidal fuzzy number, etc.), which is related to the scientificity of the solving process of problem and the reliability of the final results.

Acknowledgments The present work is supported financially by the National Natural Science Foundation of China under the grants of 51376040 and 11072057.

References

1. Li, Z.Y., Chen, C., Cao, C.X.: A prediction model for molten iron desulfuraten based on an improved RBFNN. J. Chongqing Univ. **26**(9), 119–122 (2003)
2. Zhang, H.S., Zhang, D.P., Jiang, Z.H.: Final sulfur content prediction model based on improved bp artificial neural network for hot metal pretreatment. Iron Steel **42**(3), 30–32 (2007)
3. Peng, J., Wang, S.J., Dong, Y.H.: Nitrogen content prediction in hot metal desulphurization pretreatment. J. Mater. Metall **12**(6), 255–259 (2007)
4. Zhang, H.N., Xu, A.J., Cui, J., et al.: Neural network prediction model for end point phosphorus content based on grey theory in hot metal pretreatment. Iron. Steel **47**(3), 38–43 (2013)
5. Zhou, D.K.: Injection control strategy and its realization for hot metal desulphurization with granulated magnesium. Metall. Ind. Autom. **38**(6), 23–27 (2014)
6. Tang, P., Wen, G.H., Xue, W.F., et al.: Study on optimization of desulfurization parameters of hot metal by powder injection. Res. Iron Steel **140**(5), 9–11 (2004)
7. Yu, Y.J.: Hot Metal Desulfurization Process Analysis Using Blowing CaO-Mg Powder and Parameter Optimization. Northeastern university, vol. 47 (2000)
8. Wang, N., Zhang, L.B., Zou, Z.S.: Three basic parameters determining desulphurization efficiency of hot metal by powder injection. J. Iron. Steel Res. **12**(9), 10–15 (2000)
9. Ruan, J.B., Zhang, F.S.: Factors Influencing the rate of hot metal desulphurization with granulated magnesium. J. Wuhan Univ. Sci. Technol. (Natural Science Edition) **27**(3), 242–244 (2004)
10. Guo, S.C.: Method of structuring element in fuzzy analysis(I), (II). J. Liaoning Tech. Univ. **21**(5), 670–673 (2002); **21**(6), 808–810 (2002)
11. Liu, H.T, Guo, S.C.: Fuzzy linear programming with fuzzy variables based on structured element method. Syst. Eng. Theor. Pract. **28**(6), 94–100 (2008)
12. Wang, H.D., Guo, S.C., Yue, L.Z.: An Approach to fuzzy multiple linear regression model based on the structured element theory. Syst. Eng. Theor. Pract. **34**(10), 2628–2636 (2014)
13. Chen, X.M., Chen, Q.Y.: Introduction to Metallurgical Thermodynamics. Metallurgical Industry Press. pp. 271–291 (1986)

Fuzzy Reasoning Triple I Constraint Method Based on Family of Implication Operator $L-\lambda-\Pi$

Jing-ning Shuang, Xiao-jing Hui and Jin-rui He

Abstract The new family of implication operator $L-\lambda-\Pi$ is given, which illustrates it is Łukasiewicz implication operator and Goguen implication operator of more general form. Four methods are discussed and proved based on $L-\lambda-\Pi$ implication operators, they are FMP model and FMT model fuzzy reasoning triple I constraint method and $\alpha-$triple I constraint method.

Keywords Family of implication operator $L-\lambda-\pi$ · Fuzzy reasoning · Triple I constraint method · $\alpha-$triple I constraint method

1 Introduction

In 1973, Zadeh proposed the famous composition rule of inference (CRI) method [1] to solve FMP model and FMT model, which widely used in many industrial and scientific research. Later on, Professor Wang supposed conclusions that effectively improve the CRI method, and proposed the basic ideas of universal triple I method, new triple I method, reverse triple I method and reverse triple I constraint method [2–9]. The obtained conclusions further enriched and developed the theory of fuzzy reasoning. [10, 11] discussed the universal triple I method and reverse triple I method that based on some common implication operators in fuzzy reasoning. We can find that, for different implication operator, the corresponding method and reasoning results are different, even the difference is very significant. Therefore, Wang proposed family of implication operator, that combined several implication operator into a family of implication operator. Which improved reliability of fuzzy reasoning and provided guarantees for practical application. On the basis of the above, the article constructed a new family of implication operator $L-\lambda-\Pi$, which will improve the reliability of fuzzy reasoning results.

J. Shuang (✉) · X. Hui · J. He
College of Mathematics and Computer Science, Yanan University,
Yanan 716000, People's Republic of China
e-mail: 1264905987@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_6

2 Preliminaries

Definition 1 [17] Łukasiewicz t-norm (denoted by) $\otimes_L$ is as follows:
When $a, b \in [0, 1]$, $a \otimes_L b = (a + b - 1) \vee 0$.
And an implication operator adjoin to $\otimes_L$:

$$R_L(a, b) = a \to_L b = \begin{cases} 1, a \leq b; \\ 1 - a + b, a > b. \end{cases}$$

Definition 2 [17] Goguen t-norm (denoted by) $\otimes_\Pi$ is as follows:
When $a, b \in [0, 1]$, $a \otimes_\Pi b = ab$.
And an implication operator adjoin to $\otimes_\Pi$:

$$R_\Pi(a, b) = a \to_\Pi b = \begin{cases} 1, a \leq b; \\ \frac{b}{a}, a > b. \end{cases}$$

Definition 3 (*$L - \lambda - \Pi$ implication operators*) Suppose that $\forall \lambda \in$ [0,1], the fuzzy implication operator $R(x, y) = x \to_\lambda y$ of $L - \lambda - \Pi$ implication operators is defined as follows,

$$x \to_\lambda y = \begin{cases} 1, x \leq y; \\ \frac{y+(1-\lambda)(1-x)}{1-\lambda+\lambda x}, x > y. \end{cases} \quad x, y \in [0, 1]$$

Specially, when $\lambda = 0, 1$, corresponding to R_L and R_Π implication operator, so called it $R_{L-\lambda-\Pi}$ implication operator, also called all of these implication operator is $L - \lambda - \Pi$ implication operators.

Definition 4 (*Triple I constraint principle for FMP model*) [10] Suppose that X, Y are nonempty sets, and $F(X)$, $F(Y)$ are sets of all fuzzy subsets on X, Y. When $A(x)$, $A^*(x) \in F(X)$,
$B(y) \in F(Y)$ seek out the largest $B^*(y) \in F(Y)$, so that:

$$(A(x) \to B(y)) \to (A^*(x) \to B^*(y)) \tag{1}$$

For all the $x \in X$, $y \in Y$ have the smallest possibility.

Definition 5 (*Triple I constraint principle for FMT model*) [10] Suppose that X, Y are nonempty sets, and $F(X)$, $F(Y)$are sets of all fuzzy subsets on X, Y. When $A(x) \in F(X)$, $B(y)$,
$B^*(y) \in F(Y)$, seek out the smallest $A^*(x) \in F(X)$ so that formula (1) have the smallest possibility for all the $x \in Y$, $y \in Y$.

Theorem 1 *The fuzzy implication operator $R(x, y) = x \to_\lambda y$ of $L - \lambda - \Pi$ implication operators is not reducing on the second variable.*

Proof $\forall y_1, y_2$, suppose $0 \le y_1 \le y_2 \le 1$, we need to discuss in two cases:

(1) When $x \le y_1$, that $x \le y_2$, so $R(x, y_1) = x \to y_1 = 1$, $R(x, y_2) = x \to y_2 = 1$, then $R(x, y_1) = R(x, y_2)$;
(2) When $x > y_1$, $R(x, y_1) = x \to y_1 = \frac{y_1 + (1-\lambda)(1-x)}{1-\lambda+\lambda x}$;
Meanwhile, if $x > y_2$, $R(x, y_2) = x \to y_2 = \frac{y_2 + (1-\lambda)(1-x)}{1-\lambda+\lambda x}$, then $R(x, y_1) \le R(x, y_2)$;
If $x \le y_2$, $R(x, y_2) = x \to y_2 = 1$, obviously $R(x, y_1) \le R(x, y_2)$.

Therefore, the implication operator $R(x, y) = x \to_\lambda y$ is not reducing on the second variable.

Definition 6 (*$\alpha-$triple I constraint principle for FMP model*) [10] Suppose that X, Y are nonempty sets, and $F(X)$, $F(Y)$ are sets of all fuzzy subsets on X, Y. When $A(x), A^*(x) \epsilon F(X)$,
$B(y) \epsilon F(Y)$ seek out the largest $B^*(y) \epsilon F(Y)$, so that:

$$(A(x) \to B(y)) \to (A^*(x) \to B^*(y)) \le \alpha \tag{2}$$

Definition 7 (*$\alpha-$triple I constraint principle for FMT model*) [10] Suppose that X, Y are nonempty sets, and $F(X)$, $F(Y)$ are sets of all fuzzy subsets on X, Y. When $A(x) \epsilon F(X)$,
$B(y)$, $B^*(y) \epsilon F(Y)$, seek out the smallest $A^*(x) \epsilon F(X)$ so that formula (2) have the smallest possibility for all the $x \epsilon Y$, $y \epsilon Y$.

3 The $\alpha-$Triple I Constraint Method for FMP

Remark 1 According to $\alpha-$triple I constraint principle for FMP model, formula (1) have the smallest value, that
$(A(x) \to B(y)) \to (A^*(x) \to 0) =$

$$\begin{cases} 1, A^*(x) = 0; \\ 1, A^*(x) \ne 0, R(A(x), B(y)) \le \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}; \\ \frac{\frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)} + (1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, A^*(x) \ne 0, R(A(x), B(y)) > \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}; \end{cases}$$

Based on the above analysis, First, discussed $\alpha = 1$ triple I constraint method for FMP model, Theorem 2 as follows.

Theorem 2 (Triple I constraint method for FMP model) *Suppose that X, Y are non-empty sets, When $A(x)$, $A^*(x) \epsilon F(X)$, $B(y) \epsilon F(Y)$, seek out the largest $B^*(y)$, so that formula (1) establish.*

$$B^*(y) = 1\chi_{E_y \cup K_y} + 0\chi_{(E_y \cup K_y)^C}, y \in Y$$

Which $E_y = \{x \in X | A^*(x) = 0\}$, $K_y = \{x \in X | R(A(x), B(y)) \le \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}\}$, $\chi_{E_y \cup K_y}$ *is characteristic function of* $E_y \cup K_y$, $(E_y \cup K_y)^C$ *is complementary set of* $E_y \cup K_y$.

Proof (1) When $x \in E_y \cup K_y$, according to the Remark 1, the smallest value of formula (1) is 1, then the largest $B^*(y) = 1$;

(2) When $x \in (E_y \cup K_y)^C$, that $A^*(x) \neq 0, R(A(x), B(y)) > \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}$;, according to the Remark 1, the smallest value of formula (1) is $\frac{\frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}$, then the largest $B^*(y) = 0$;

If $\exists y_0 \in Y$, satisfy $D(y_0) > B^*(y_0) = 0$, then $D(y_0)$ must not make formula (1) establish. Because of $D(y_0) > B^*(y_0)$, so we need to discuss in two cases:
When $A^*(x_0) \leqslant D(y_0)$,
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) = R(A(x_0), B(y_0)) \to 1 = 1 > \frac{\frac{(1-\lambda)(1-A^*(x_0))}{1-\lambda+\lambda A^*(x_0)}+(1-\lambda)(1-R(A(x_0),B(y_0)))}{1-\lambda+\lambda R(A(x_0),B(y_0))}$; When $A^*(x_0) > D(y_0)$,
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) = R(A(x_0), B(y_0)) \to \frac{D(y_0)+(1-\lambda)(1-A^*(x_0))}{1-\lambda+\lambda A^*(x_0)}$, so we need to discuss in two cases again:
If $R(A(x_0), B(y_0)) \le \frac{D(y_0)+(1-\lambda)(1-A^*(x_0))}{1-\lambda+\lambda A^*(x_0)}$, that
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) = 1$;
If $R(A(x_0), B(y_0)) > \frac{D(y_0)+(1-\lambda)(1-A^*(x_0))}{1-\lambda+\lambda A^*(x_0)}$, that
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) > \frac{\frac{(1-\lambda)(1-A^*(x_0))}{1-\lambda+\lambda A^*(x_0)}+(1-\lambda)(1-R(A(x_0),B(y_0)))}{1-\lambda+\lambda R(A(x_0),B(y_0))}$;
Therefore, when $x \in (E_y \cup K_y)^C$, $B^*(y)$ have the largest value 0.
To sum up, the Theorem 2 is a constant establishment.

Remark 2 According to the discussion of Remark 1, when $A^*(x) = 0$ or when $A^*(x) \neq 0$ and $R(A(x), B(y)) \le \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}$, the smallest value of formula (1) is 1. Meanwhile, if $\alpha = 1$, then the largest $B^*(y)$ of formula (2) is also 1; if $\alpha < 1$, formula (2) is not establishment. When $A^*(x) \neq 0, R(A(x), B(y)) > \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}$, if $\alpha = 1$, the largest $B^*(y)$ of formula (2) is also 1; When $\alpha < \frac{\frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}$, formula (2) is not establishment. Therefore, formula (2) need to establish, the value rang of α must satisfy $\alpha \in [\frac{\frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, 1]$. Next we only discussed when $\alpha \in (\frac{\frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, 1)$.

Theorem 3 (α-triple I constraint method for FMP model) *Suppose that X, Y are nonempty sets, When* $A(x)$, $A^*(x) \epsilon F(X)$, $B(y) \epsilon F(Y)$, *seek out the largest* $B^*(y)$, *so that formula (2) establish. When* $x \in E_y$,

$$B^*(y) = \inf\{A^*(x) \wedge [(1-\lambda+\lambda A^*(x))(R(A(x), B(y)) - 1) + A^*(x)] \wedge [A^*(x) - (1-\alpha)$$

$(1-\lambda+\lambda R(A(x), B(y)))(1-\lambda+\lambda A^*(x))-(1-R(A(x), B(y)))(1-\lambda+\lambda A^*(x))]\}$,
$y \in Y$
Which $E_y = \{x \in X | R(A(x), B(y)) > \frac{(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}\}$.

Proof On the one hand, $\forall y \in Y$, $C(y) \in F(Y)$, satisfy $C(y) < B^*(y)$, then $C(y)$ make the formula (2) establish. In fact, because of $C(y) < A^*(x)$ and $C(y) < (1-\lambda+\lambda A^*(x))(R(A(x), B(y))-1)+A^*(x)$ and $C(y) < A^*(x)-(1-\alpha)(1-\lambda+\lambda R(A(x), B(y)))(1-\lambda+\lambda A^*(x))-(1-R(A(x), B(y)))(1-\lambda+\lambda A^*(x))$, then
$(A(x) \to B(y)) \to (A^*(x) \to C(y)) = \frac{\frac{C(y)+(1-\lambda)(1-A^*(x))}{1-\lambda+\lambda A^*(x)}+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))} < \alpha$;
Therefore $(A(x) \to B(y)) \to (A^*(x) \to C(y)) < \alpha$, then $C(y)$ make the formula (2) establish. On the other hand, if $\exists y_0 \in Y$, satisfy $D(y_0) \geq B^*(y_0)$, then $D(y_0)$ isnt make the formula (2) establish. In fact, because of $D(y_0) \geq B^*(y_0)$,
$D(y_0) \geq A^*(x_0) \wedge [(1-\lambda+\lambda A^*(x_0))(R(A(x_0), B(y_0))-1)+A^*(x_0)] \wedge [A^*(x_0)-(1-\alpha)(1-\lambda+\lambda R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))-(1-R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))]$,
so we need to discuss in two cases:
(1) When $D(y_0) \geq A^*(x_0)$, that
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) = R(A(x_0), B(y_0)) \to 1 = 1 \geq \alpha$;
(2) When $D(y_0) < A^*(x_0)$, then $D(y_0) \geq [(1-\lambda+\lambda A^*(x_0))(R(A(x_0), B(y_0))-1)+A^*(x_0)] \wedge [A^*(x_0)-(1-\alpha)(1-\lambda+\lambda R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))-(1-R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))]$, so we need to discuss in two cases again:
If $D(y_0) \geq (1-\lambda+\lambda A^*(x_0))(R(A(x_0), B(y_0))-1)+A^*(x_0)$, then
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) = 1$;
If $D(y_0) < (1-\lambda+\lambda A^*(x_0))(R(A(x_0), B(y_0))-1)+A^*(x_0)$, $D(y_0) \geq A^*(x_0)-(1-\alpha)(1-\lambda+\lambda R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))-(1-R(A(x_0), B(y_0)))(1-\lambda+\lambda A^*(x_0))$, then
$(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) > \alpha$;
Therefore $(A(x_0) \to B(y_0)) \to (A^*(x_0) \to D(y_0)) > \alpha$, then $D(y_0)$ isnt make the formula (2) establish.
To sum up, the Theorem 3 is a constant establishment.

4 The $\alpha-$Triple I Constraint Method for FMT

Remark 3 According to triple I constraint principle for FMT model $A^*(x) = 1$, formula (1) have the smallest value, that
$(A(x) \to B(y)) \to (1 \to (B^*(y)) =$

$$\begin{cases} 1, R(A(x), B(y)) \leq B^*(y); \\ \frac{B^*(y)+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, R(A(x), B(y)) > B^*(y); \end{cases}$$

Based on the above discussion, when $\alpha = 1$, triple I constraint method for FMT model as follows.

Theorem 4 (Triple I constraint method for FMT model) *Suppose that X, Y are nonempty sets, When $A(x) \epsilon F(X)$, $B(y)$, $B^*(y) \epsilon F(Y)$, seek out the smallest $A^*(x)$, so that formula (1) establish.*

$$A^*(x) = 1\chi_{E_x} + 0\chi_{(E_x)^C}, x \in X$$

Which $E_x = \{y \in Y | R(A(x), B(y)) > B^*(y)\}$.

Proof (1) If $y \in E_x$, $R(A(x), B(y)) > B^*(y)$;
According to the Remark 3, when $R(A(x), B(y)) > B^*(y)$, the smallest value of formula (1) is $\frac{B^*(y)+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}$, then the smallest $A^*(x) = 1$;
If $\exists x_0 \in X$, make $C(x_0) < A^*(x_0) = 1$, we need to discuss in two cases:
If $C(x_0) \le B^*(y_0)$, that
$(A(x_0) \to B(y_0)) \to (C(x_0) \to B^*(y_0)) = R(A(x_0), B(y_0)) \to 1 = 1$;
If $C(x_0) > B^*(y_0)$, that
$(A(x_0) \to B(y_0)) \to (C(x_0) \to B^*(y_0)) = R(A(x_0), B(y_0)) \to \frac{B^*(y_0)+(1-\lambda)(1-C(x_0))}{1-\lambda+\lambda C(x_0)}$, then we need to discuss in two cases again:
When $R(A(x_0), B(y_0)) \le \frac{B^*(y_0)+(1-\lambda)(1-C(x_0)}{1-\lambda+\lambda C(x_0)}$, then
$(A(x_0) \to B(y_0)) \to (C(x_0) \to B^*(y_0)) = 1$;
When $R(A(x_0), B(y_0)) > \frac{B^*(y_0)+(1-\lambda)(1-C(x_0))}{1-\lambda+\lambda C(x_0)}$, then $(A(x_0) \to B(y_0)) \to (C(x_0) \to B^*(y_0)) > \frac{B^*(y_0)+(1-\lambda)(1-R(A(x_0),B(y_0)))}{1-\lambda+\lambda R(A(x_0),B(y_0))}$;
Therefore, when $y \in E_x$, $A^*(x)$ have the smallest value 1.
(2) When $y \in (E_x)^C$, $R(A(x), B(y)) \le B^*(y)$;
According to the Remark 3, when $R(A(x), B(y)) \le B^*(y)$, the smallest value of formula (1) is 1, then the smallest $A^*(x) = 1$;
To sum up, the Theorem 4 is a constant establishment.

Remark 4 According to the discussion of Remark 3, when $R(A(x), B(y)) \le B^*(y)$, the smallest value of formula (1) is 1. Meanwhile, if $\alpha = 1$, the smallest $A^*(x) = 0$; If $\alpha < 1$, formula (2) is not establishment. When $R(A(x), B(y)) > B^*(y)$, if $\alpha = 1$, the smallest $A^*(x) = 0$; If $\alpha < \frac{B^*(y)+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}$, formula (2) is not establishment. Therefore, formula (2) need to establish, the value rang of α must satisfy $\alpha \in [\frac{B^*(y)+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, 1]$, Next we only discussed when $\alpha \in (\frac{B^*(y)+(1-\lambda)(1-R(A(x),B(y)))}{1-\lambda+\lambda R(A(x),B(y))}, 1)$.

Theorem 5 ($\alpha-$triple I constraint method for FMT model) *Suppose that X, Y are nonempty sets, When $A(x) \epsilon F(X)$, $B(y)$, $B^*(y) \epsilon F(Y)$, seek out the smallest $A^*(x)$ so that formula (2) establish. When $y \in E_x$,*
$A^*(x) = \sup\{B^*(y) \vee [\frac{B^*(y)-R(A(x),B(y))}{1-\lambda+\lambda R(A(x),B(y))} + 1] \vee \frac{B^*(y)+(1-\lambda)[1-\alpha R(A(x),B(y))+(1-\alpha)(1-\lambda)(1-R(A(x),B(y)))]}{1-\lambda+\alpha\lambda R(A(x),B(y))-\lambda(1-\alpha)(1-\lambda)(1-R(A(x),B(y)))}\}$, $x \in X$;
Which $E_x = \{y \in Y | R(A(x), B(y)) > B^*(y)\}$.

Proof On the one hand, $\forall x \in X$, $C(x) \in F(X)$, satisfy $C(x) > A^*(x)$, then $C(x)$ make the formula (2) establish. In fact, because of $C(x) > A^*(x)$, $C(x) > B^*(y) \vee [\frac{B^*(y)-R(A(x),B(y))}{1-\lambda+\lambda R(A(x),B(y))} + 1] \vee \frac{B^*(y)+(1-\lambda)[1-\alpha R(A(x),B(y))+(1-\alpha)(1-\lambda)(1-R(A(x),B(y)))]}{1-\lambda+\alpha\lambda R(A(x),B(y))-\lambda(1-\alpha)(1-\lambda)(1-R(A(x),B(y)))}$, then $(A(x) \to B(y)) \to (C(x) \to B^*(y)) = R(A(x), B(y)) \to \frac{B^*(y)+(1-\lambda)(1-C(x))}{1-\lambda+\lambda C(x)} < \alpha$;
On the other hand, if $\exists x_0 \in X$, satisfy $D(x_0) \leq A^*(x_0)$, then $D(x_0)$ isnt make the formula (2) establish. In fact, because of $D(x_0) \leq A^*(x_0)$, $y_0 \in Y$, make $D(x_0) \leq B^*(y_0) \vee [\frac{B^*(y_0)-R(A(x_0),B(y_0))}{1-\lambda+\lambda R(A(x_0),B(y_0))} + 1] \vee \frac{B^*(y_0)+(1-\lambda)[1-\alpha R(A(x_0),B(y_0))+(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))]}{1-\lambda+\alpha\lambda R(A(x_0),B(y_0))-\lambda(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))}$,
we need to discuss in two cases:
(1) When $D(x_0) \leq B^*(y_0)$, then
$(A(x_0) \to B(y_0)) \to (D(x_0) \to B^*(y_0)) = R(A(x_0), B(y_0)) \to 1 = 1$;
(2) When $D(x_0) > B^*(y_0)$, then
$D(x_0) \leq [\frac{B^*(y_0)-R(A(x_0),B(y_0))}{1-\lambda+\lambda R(A(x_0),B(y_0))} + 1] \vee \frac{B^*(y_0)+(1-\lambda)[1-\alpha R(A(x_0),B(y_0))+(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))]}{1-\lambda+\alpha\lambda R(A(x_0),B(y_0))-\lambda(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))}$,
we need to discuss in two cases again:
If $D(x_0) \leq \frac{B^*(y_0)-R(A(x_0),B(y_0))}{1-\lambda+\lambda R(A(x_0),B(y_0))} + 1$, then
$(A(x_0) \to B(y_0)) \to (D(x_0) \to B^*(y_0)) = 1 > \alpha$;
If $D(x_0) > \frac{B^*(y_0)-R(A(x_0),B(y_0))}{1-\lambda+\lambda R(A(x_0),B(y_0))} + 1$, that
$D(x_0) \leq \frac{B^*(y_0)+(1-\lambda)[1-\alpha R(A(x_0),B(y_0))+(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))]}{1-\lambda+\alpha\lambda R(A(x_0),B(y_0))-\lambda(1-\alpha)(1-\lambda)(1-R(A(x_0),B(y_0)))}$, then
$(A(x_0) \to B(y_0)) \to (D(x_0) \to B^*(y_0)) > \alpha$.
To sum up, the Theorem 5 is a constant establishment.

5 Conclusion

This paper constructs a family of implication operation $L-\lambda-\Pi$, which improve the reliability of fuzzy reasoning results. And four methods are discussed and proved that based on $L-\lambda-\Pi$ implication operators, they are FMP model and FMT model fuzzy reasoning triple I constraint method and $\alpha-$triple I constraint method. The obtained conclusions are enriched and developed the relevant theory of fuzzy reasoning triple I constraint method, and reduce adventure in the practical application. In other papers, based on the new family of implication operation $L-\lambda-\Pi$, we discuss FMP model and FMT model fuzzy reasoning reverse triple I method, reverse triple I constraint method.

Acknowledgments The National Natural Science Foundation of China under Grant No. 11471007; The National Science Foundation of Shanxi Province of China under Grant No. 2014JM1020.

References

1. Zadeh, L.A.: Outline of new approach to the analysis of complex systems and decision processes [J]. IEEE Trans. Syst. Man Cybern. **3**(1), 28–33 (1973)
2. Wang, G.: Total complication triple I method in fuzzy reasoning[J]. Sci. China: Ser E **29**(1), 43–53 (1999)
3. Wang, G.: A new method for fuzzy reasoning[J]. Fuzzy Syst. Math. **13**(3), 1–9 (1999)
4. Wang, G., Song, Q.: A new triple I method and its basement of logic[J]. Prog. Nat. Sci. **13**(6), 575–581 (2003)
5. Pei, D.: Two triple I methods and their consistency of FMP problem[J]. Fuzzy Syst. Math. **15**(4), 1–6 (2001)
6. Pei, D.: Full implication triple I algorithms and their consistency in fuzzy reasoning[J]. J. Math. Res. Exposition **24**(2), 359–368 (2004)
7. Song, S., Feng, C., Wu, C.: Based on full implication triple I constraint method in fuzzy reasoning [J]. Prog. Nat. Sci. **10**(10), 884–889 (2000)
8. Song, S., Wu, C.: Fuzzy reasoning reverse triple I method[J]. Sci. China: Ser E **32**(2), 230–246 (2002)
9. Song, S., Wu, C.: Fuzzy reasoning reverse triple I constraint method[J]. Prog. Nat. Sci. **12**(1), 95–100 (2002)
10. Peng, J., Hou, J., Li, H.: Reverse triple I method under commonly used fuzzy implication operator[J]. Prog. Nat. Sci. **15**(4), 404–410 (2005)
11. Peng, J.: Fuzzy reasoning total complication triple I constraint method under commonly used fuzzy implication operator[J]. Prog. Nat. Sci. **15**(5), 539–546 (2005)
12. Mingwang, W.: Generalized tautologies in parametric Kleenes systems[J]. Fuzzy Syst. Math. **14**(1), 1–7 (2000)
13. Wu, H: Generalized tautologies in modefiable Kleenes systems[J]. Sci. China: Ser E **32**(2), 224–229 (2002)
14. Wang, G., Rong, L.: Generalized tautologies of the Systems Ha[J]. J. Shanxi Normal Univ.: Nat. Sci. Ed. **31**(2), 1–11 (2003)
15. Zhang, X., Meng, G.: Families of implication operators and their application[J]. Chin. J. Comput. **30**(3), 448–453 (2007)
16. Wang, G.: Introduction to Mathematical Logic and Resolution Principle. Science in China Press, Beijing (2006) (in Chinese)
17. Wang, G.: Non-classical Mathematical Logic and Approximate. Science in China Press, Beijing (2008) (in Chinese)

Fuzzy Information Fusion Approach for Supplier Selection

Guangxu Li, Gang Kou and Yi Peng

Abstract Supplier selection plays an important role in supply chain system. In order to select the suitable suppliers, some methods of supplier selection have been studied extensively in fuzzy environment. In this paper, a fuzzy information fusion approach based on generalized fuzzy numbers (GFNs) is proposed to select the best supplier. Some aggregation operators with GFNs are also proposed. In addition, a new ranking formula based on mean values of GFNs is adopted to rank the suppliers. Finally, an empirical study of supplier selection is introduced to illustrate the proposed method. The results indicate that the proposed method could meet the different evaluation requirements of decision makers and offer an effective and practical way to select the best supplier.

Keywords Information fusion · Supplier selection · Fuzzy decision making · Generalized fuzzy numbers · Mean value

1 Introduction

The supplier selection in its nature is a multi-criteria decision-making (MCDM) problem since some conflicting criteria have influence on evaluation and selection of suppliers [1]. MCDM methods are important issues for decision science, and they have been applied many fields [2–4]. S Moreover, MCDM methods have been also applied to evaluate or rank the suppliers with respect to multiple attributes [5, 6]. However, some supplier selection problems usually involve in uncertain, indefinite, imprecise and subjective data.

G. Li · Y. Peng
School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, China

G. Kou (✉)
School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611731, Sichuan, China
e-mail: kougang@swufe.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_7

Fuzzy set theory, initially proposed by Zadeh [7], has been extensively applied to objectively reflect the ambiguities in human judgment. One of applications of the fuzzy theory is used in fuzzy MCDM. So some fuzzy MCDM methods were proposed to evaluate the supplier selection. Chen et al. [8] proposed a hierarchy MCDM model based on fuzzy-sets theory to deal with the supplier selection problems in the supply chain system. Junior et al. [9] presents a comparative analysis of fuzzy AHP and fuzzy TOPSIS methods in the context of supplier selection decision making. Senvar et al. [10] proposed a fuzzy PROMETHEE method to select suppliers. However, they considered the linear fuzzy numbers and the viewpoints of experts cannot capture the sophisticated nuances when they selected suppliers. In order to meet the different evaluation requirements of decision makers, Li et al. [11] proposed a new form of fuzzy number, named a generalized fuzzy number (GFN). The GFN is a nonlinear fuzzy number, and it can represent different information when the parameter changes in supplier selection. Therefore, in the paper, GFNs are applied to express the values of attributes in supplier selection.

The supplier selection in its nature is a multi-criteria decision-making (MCDM) problem. Wu et al. [12] proposed a fuzzy multi-objective programming model to decide on supplier selection taking risk factors into consideration. Lin [13] proposed to adopt the fuzzy analytic network process (FANP) approach first to identify top suppliers by considering the effects of interdependence among selection criteria and to handle inconsistent and uncertain judgments. Nazari-Shirkouhi et al. [14] presented a two-phase fuzzy multi-objective linear programming to solve supplier selection and order allocation problem. However, these methods did not consider the multiple persons decision making processes. In order to integrate the expert information, the uniform information are usually aggregated by aggregation operators. For example, an ordered weighted averaging (OWA) aggregation operator was applied in [15]. Chiclana et al. [16] designed a new choice scheme using the concept of fuzzy majority and a new aggregation operator, which is called ordered weighted geometric (OWG) operator. Therefore, some other aggregation operators have been developed to aggregate preference information. Such as: an uncertain OWA operator [17] and a generalized ordered weighted logarithm aggregation (OWLA) operator [18]. However, these aggregation operators are linear operators, and do not capture the sophisticated nuances in the aggregated value. Moreover, the interrelations among the integrated preferences are not considered in the aggregation processes. In order to overcome the shortcomings of the aggregation operators, the power average operator was proposed by Yager [19]. Xu and Yager [20] developed a power geometric operator and gave the applications to MCDM. However, aggregation operators with GFNs are rarely researched. Based on this idea, we propose some nonlinear aggregation operators of generalized fuzzy numbers (GFNs), and give the applications in supplier selection.

In fuzzy MCDDM, many ranking methods have been done on how to rank the alternatives based on the priority vector. For example, Ergu et al. [21] advanced a rapid decision making method to assess the key factors of risks. Asady [22] gave a revised ranking method based on deviation degree. Matarazzo and Munda [23] presented a ranking fuzzy numbers method based on triangular fuzzy numbers. But they did not consider the mean values of GFNs. Based on the mean values of fuzzy numbers, a ranking formula based on the mean values of GFNs is proposed to apply in supplier selection in our paper.

2 Definitions and Properties of GFNs

Definition 1 [11] A GFN $\tilde{A}$ is given by $\tilde{A} = (a, b, c, d)_n, n > 0, 0 \leq a \leq b \leq c \leq d$, if the membership function $\mu_{\tilde{A}}(x) : R \rightarrow [0, 1]$ is defined as follows

$$\tilde{A} = \begin{cases} (\frac{x-a}{b-a})^n, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ (\frac{d-x}{d-c})^n, & c \leq x \leq d \\ 0, & others \end{cases}$$

In order to intuitively illustrate the GFN, we suppose $a = 2, b = 6, c = 8, d = 12$, and give the graphs of the GFN with different values of n in Fig. 1.

From the graphs of the GFN, some aspects of the GFNs are proposed. For example, trapezoidal fuzzy numbers or triangular fuzzy numbers are the special forms of GFNs. Let $n = 1$, then $\tilde{A}$ is reduced to a trapezoidal fuzzy number, and if $n = 1$ and $b = c$, then $\tilde{A}$ is reduced to a triangular fuzzy number. Moreover, if $n > 1$, then the left and right branches contract for the membership function of the GFN, and if $0 < n < 1$, then the left and right branches expand for the membership function of the GFN. With diversification of the parameter n, some characteristics about the GFNs are also given. For example, with the values of parameter n increasing, the fuzzy degrees of $\tilde{A}$ will be reduced, and then the information of $\tilde{A}$ will be more

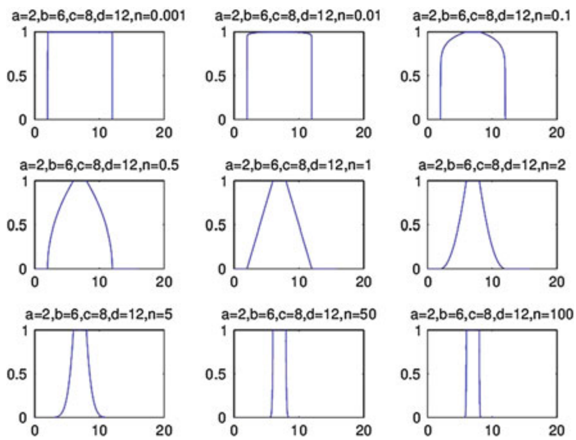


Fig. 1 GFN with different values of n

clearly. Thus, the GFN is applied in the decision making process, the subjectivity of selecting the model parameters should be reduced, and the robustness of the model is improved. Therefore, the GFN could be suitable for all kinds of fuzzy environment. Moreover, some main operations of GFNs are also given by [11].

3 Nonlinear Aggregation Operators with GFNs

Information aggregation is an important process in GDM. PA operator, introduced by Yager [19], is a nonlinear operator which allows argument values to support each other in the aggregation process. The definition of PA operator is given as follows

Definition 2 [19] Let $a_1, a_2, \ldots, a_n$ be a collection of arguments, then the power average (PA) operator $PA(a_1, a_2, \ldots, a_n)$ is defined as

$$PA(a_1, a_2, \ldots, a_n) = \frac{\sum_{i=1}^{n}(1 + T(a_i))a_i}{\sum_{i=1}^{n}(1 + T(a_i))} \tag{1}$$

where

$$T(a_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} Sup(a_i, a_j) \tag{2}$$

$Sup(a, b)$ is denoted as the support between a and b. Three properties of the support are given as follows: (1) $Sup(a, b) \in [0, 1]$; (2) $Sup(a, b) = Sup(b, a)$; (3) $Sup(a, b) \geq Sup(x, y)$ if $|a - b| \leq |x - y|$.

Obviously, the support measure is essentially a similarity index. The more similar, the closer the two values, and the more they support each other.

Based on Definition 2, some nonlinear aggregation operators with GFNs can be given as follows

Definition 3 Let $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n$ be a collection of GFNs, then the power average (PA) operator of GNFs is defined as

$$GFNsPA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \frac{\sum_{i=1}^{n}(1 + T(\tilde{A}_i))\tilde{A}_i}{\sum_{i=1}^{n}(1 + T(\tilde{A}_i))} \tag{3}$$

where

$$T(\tilde{A}_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} Sup(\tilde{A}_i, \tilde{A}_j) \tag{4}$$

$Sup(\tilde{A}, \tilde{B})$ is denoted as the support between $\tilde{A}$ and $\tilde{B}$. Three properties of the support are also given as follows: (1) $Sup(\tilde{A}, \tilde{B}) \in [0, 1]$; (2) $Sup(\tilde{A}, \tilde{B}) = Sup(\tilde{B}, \tilde{A})$; (3) If

$D_H(\tilde{A},\tilde{B}) \le D_H(\tilde{X},\tilde{Y})$, then $Sup(\tilde{A},\tilde{B}) \ge Sup(x,y)$, and the formula of D_H is given by [11, 17].

Obviously, the more similar, the closer two values, the more they support each other. Based on the properties of the supports, the formula of support can be given as follows

$$Sup(\tilde{A}_i,\tilde{A}_j) = 1 - \frac{D_H(\tilde{A}_i,\tilde{A}_j)}{\sum_{\substack{j=1 \\ j\neq i}}^{n} D_H(\tilde{A}_i,\tilde{A}_j)} \tag{5}$$

Especially, if $D_H(\tilde{A}_i,\tilde{A}_j) = 0$, then we stipulate $Sup(\tilde{A}_i,\tilde{A}_j) = 1$.

The PA operator of GFNs exhibits a number of properties desirable for an aggregation operator.

(1) Commutativity: If $(\tilde{C}_1,\tilde{C}_2,\ldots,\tilde{C}_n)$ is any permutation of $(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n)$, then $GFNsPA(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) = GFNsPA(\tilde{C}_1,\tilde{C}_2,\ldots,\tilde{C}_n)$.
(2) Idempotency: If $\tilde{A}_i = \tilde{A}_j = \tilde{A}(i \neq j)$, then $GFNsPA(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) = \tilde{A}$.
(3) Boundedness: The PA operator of GFNs satisfies $min(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) \le GFNsPA(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) \le max(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n)$.
(4) Generality: If $Sup(\tilde{A}_i,\tilde{A}_j) = k$, for all $i \neq j$, then the PA operator of GFNs satisfies $GFNsPA(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) = \sum_{i=i}^{n} \frac{\tilde{A}_i}{n}$.

Proof (1) Let $(\tilde{C}_1,\tilde{C}_2,\ldots,\tilde{C}_n)$ is any permutation of $(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n)$, $\tau : \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ be a permutation function, there exists one and only one $\tilde{C}_j$ such that $\tilde{A}_i = \tilde{A}_{\tau(i)} = \tilde{C}_j$, $i = 1,2,\ldots,n$, then we have

$$T(\tilde{A}_i) = \sum_{\substack{j=1 \\ j\neq i}}^{n} Sup(\tilde{A}_i,\tilde{A}_j) = \sum_{\substack{j=1 \\ j\neq i}}^{n} Sup(\tilde{A}_{\tau(i)},\tilde{A}_{\tau(i)}) = \sum_{\substack{j=1 \\ j\neq i}}^{n} Sup(\tilde{C}_j,\tilde{C}_i) = T(\tilde{A}_j)$$

By formulas (7)–(9), we obtain

$$\begin{aligned} GFNsPA(\tilde{C}_1,\tilde{C}_2,\ldots,\tilde{C}_n) &= \sum_{i=1}^{n}(1+T(\tilde{C}_j))\tilde{C}_j / \sum_{i=1}^{n}(1+T(\tilde{C}_j)) \\ &= \sum_{i=1}^{n}(1+T(\tilde{A}_i))\tilde{A}_i / \sum_{i=1}^{n}(1+T(\tilde{A}_i)) \\ &= GFNsPA(\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n) \end{aligned}$$

Therefore, the property 1 holds.

(2) If $\tilde{A}_i = \tilde{A}_j = \tilde{A}(i \neq j)$, then $D_H(\tilde{A}_i, \tilde{A}_j) = 0$ and $Sup(\tilde{A}_i, \tilde{A}_j) = 1$, so we can get

$$\begin{aligned} GFNsPA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) &= \sum_{i=1}^{n} (1 + (n-1))\tilde{A}_i / \sum_{i=1}^{n} (1 + (n-1)) \\ &= \sum_{i=1}^{n} n\tilde{A}_i / \sum_{i=1}^{n} n = \tilde{A} \end{aligned}$$

Therefore, the property 2 holds.

(3) Let $min(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \tilde{A}_{min}$ and $max(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \tilde{A}_{max}$, then we have

$$\tilde{A}_{min} \sum_{i=1}^{n} (1 + T(\tilde{A}_i)) \leq \sum_{i=1}^{n} (1 + T(\tilde{A}_i))\tilde{A}_i \leq \tilde{A}_{max} \sum_{i=1}^{n} (1 + T(\tilde{A}_i))$$

since $\sum_{i=1}^{n} (1 + T(\tilde{A}_i)) > 0$, again we get

$$\tilde{A}_{min} \leq \sum_{i=1}^{n} (1 + T(\tilde{A}_i))\tilde{A}_i / \sum_{i=1}^{n} (1 + T(\tilde{A}_i)) \leq \tilde{A}_{max}$$

and then, we obtain

$$min(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) \leq GFNsPA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) \leq max(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n)$$

Therefore, the property 3 holds.

(4) If $Sup(\tilde{A}_i, \tilde{A}_j) = k$, for all $i \neq j$, then $T(\tilde{A}_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} Sup(\tilde{A}_i, \tilde{A}_j) = (n-1)k$, thus we get

$$GFNsPA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \sum_{i=1}^{n} (1 + (n-1)k)\tilde{A}_i / \sum_{i=1}^{n} (1 + (n-1)k) = \sum_{i=i}^{n} \frac{\tilde{A}_i}{n}$$

This means when all the matrix similarity are not different, the PA operator of GFNs is simplified as the arithmetic average operator of GFNs. Therefore, the property 4 holds.

Definition 4 The weighted power average (WPA) operator of GFNs is defined as

$$GFNsWPA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \frac{\sum_{i=1}^{n} (1 + T(\tilde{A}_i))\tilde{A}_i \omega_i}{\sum_{i=1}^{n} (1 + T(\tilde{A}_i))\omega_i} \tag{6}$$

where

$$T(\tilde{A}_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} \omega_i Sup(\tilde{A}_i, \tilde{A}_j) \tag{7}$$

$Sup(\tilde{A}, \tilde{B})$ is denoted as the support between $\tilde{A}$ and $\tilde{B}$.

Similarly, the WPA operator of GFNs has the properties such as idempotency, boundedness, generality, but commutativity property does not hold.

Definition 5 Let $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n$ be a collection of GFNs, then the power geometric average (PGA) operator of GNFs is defined as

$$GFNsPGA(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n) = \prod_{i=1}^{n} \tilde{A}_i^{\frac{(1+T(\tilde{A}_i))}{\sum_{i=1}^{n}(1+T(\tilde{A}_i))}} \tag{8}$$

where

$$T(\tilde{A}_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} Sup(\tilde{A}_i, \tilde{A}_j), Sup(\tilde{A}_i, \tilde{A}_j) = 1 - \frac{D_H(\tilde{A}_i, \tilde{A}_j)}{\sum_{\substack{j=1 \\ j \neq i}}^{n} D_H(\tilde{A}_i, \tilde{A}_j)}.$$

The PGA operator of GFNs also exhibits a number of properties such as commutativity, idempotency, boundedness and generality.

After the information are integrated by the nonlinear operators of GFNs, the results in the aggregation process are also showed by GFNs. In order to evaluate the results, we should rank the fuzzy numbers. A ranking formula based on mean values of GFNs is given in the next section.

4 Ranking Formula in GDM Based on Mean Values of GFNs

In the section, a ranking formula based on mean values of GFNs is proposed. Firstly, some definitions of mean values of fuzzy numbers should be given as follows.

Definition 6 [23] Let $\tilde{A} = (a, b, c)$ be a triangular fuzzy number, $\mu_{\tilde{A}}(x)$ be the membership function, and x be an element of the triangular fuzzy number $\tilde{A}$, the mean value of $\tilde{A}$ is given by

$$E(\tilde{A}) = \frac{a+b+c}{3} \tag{9}$$

Based on Definition 6, the definition of mean value with a GFN is given as follows

Definition 7 Let $\tilde{A} = (a, b, c, d)_n, n > 0, 0 \le a \le b \le c \le d$ be a GFN, $\mu_{\tilde{A}}(x)$ be the membership function, and x be an element of the GFN $\tilde{A}$, the mean value of GFN is given by

$$E(\tilde{A}) = \frac{\int_{-\infty}^{+\infty} x\mu_{\tilde{A}}(x)dx}{\int_{-\infty}^{+\infty} \mu_{\tilde{A}}(x)dx} = \frac{f(n)}{g(n)} \tag{10}$$

where

$$f(n) = \frac{c^2 - b^2}{2}n^2 + (\frac{c^2 - b^2}{2} + cd - ab)n + d^2 - a^2 \tag{11}$$

$$g(n) = (c - b)n^2 + (d - a + 2c - 2b)n + 2(d - a) \tag{12}$$

Especially, if $n = 1$, $\tilde{A}$ is reduced to a positive trapezoidal fuzzy number, and $f(1) = c^2 + d^2 - b^2 - a^2 + cd - ab$ and $g(1) = 3(c + d - a - b)$, then the mean value of a trapezoidal fuzzy number is given as follows

$$E(\tilde{A}) = \frac{c^2 + d^2 - b^2 - a^2 + cd - ab}{3(c + d - a - b)}$$

If $n = 1$ and $b = c$, $\tilde{A}$ is reduced to a positive triangular fuzzy number, and $f(1) = bd - ab + d^2 - a^2$ and $g(1) = 3(d - a)$, then we have

$$E(\tilde{A}) = \frac{bd - ab + d^2 - a^2}{3(d - a)} = \frac{(d - a)(a + b + d)}{3(d - a)} = \frac{a + b + d}{3}$$

The results are consistent with the Definition 6. For convenience, the alternatives could be expressed as $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ and the evaluation attributes could be expressed as $\tilde{c} = (\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_n)$. The attributes are additively independent. $\tilde{x}_{ij}$ is the assessed value of attribute $\tilde{c}_j$ of alternative $\tilde{x}_i$, and is expressed in a GFN in this paper. The different values of $\tilde{x}_{ij}$ can be represented by a matrix $\tilde{V} = (\tilde{x}_{ij})_{mn}$, which is called the decision making matrix. The vector of attribute weights is $\omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. To eliminate the difference of the attribute indexes on the dimension, each attribute index is normalized

$$\tilde{r}_{ij} = \begin{cases} \frac{\tilde{x}_{ij}}{x_j^+} \wedge 1, \ \forall i \in M, j \in I_1 \\ \frac{x_j^-}{\tilde{x}_{ij}} \wedge 1, \ \forall i \in M, j \in I_2 \end{cases}$$

where I_1 is associated with a set of benefit criteria, and I_2 is associated with a set of cost criteria and $M = \{1, 2, \ldots, m\}$.

When the information are integrated by nonlinear operators of GFNs, the results in the aggregation process are also showed by GFNs. In order to evaluate the results, a five-step hybrid ranking formula based on mean values of GFNs is given as follows:

Step 1. Set up decision matrix based on GFNs and normalize decision-making matrix

Step 2. Integrate information by nonlinear operators of GFNs based on formulas (5)–(7)

Step 3. Calculate the comprehensive property values by weighted arithmetic average

Step 4. Calculate the mean values of GFNs based on formulas (10)–(12)

Step 5. Rank the alternatives based on the mean values of GFNs.

5 Numerical Example

A numerical example is considered in the section. Assume that a car company needs to purchase some auto parts from three auto parts suppliers A_1, A_2, A_3. Four attributes: C_1 (the producing ability), C_2 (the ability of capital currency), C_3 (the ability of research), C_4 (the level of technology innovation) are taken into consideration and the weight vector is given by $\omega = (0.2, 0.2, 0.3, 0.3)^T$. During the decision-making process, three experts e_1, e_2, e_3 are invented to evaluate the three auto parts suppliers. Their weighting vector is given by $\lambda = (0.3, 0.3, 0.4)^T$. After analyzing and assessing these suppliers, the decision making process is shown as follows

Step 1. Set up group decision matrix based on GFNs and normalize the decision matrix in Tables 1, 2, 3, 4, 5 and 6, respectively.

Table 1 Decision matrix given by expert e_1

	A_1	A_2	A_3
C_1	$(1, 2, 3, 4)_n$	$(4, 5, 6, 7)_n$	$(2, 4, 5, 8)_n$
C_2	$(2, 3, 4, 5)_n$	$(1, 3, 5, 6)_n$	$(2, 3, 4, 5)_n$
C_3	$(3, 4, 6, 7)_n$	$(4, 6, 7, 9)_n$	$(1, 3, 6, 7)_n$
C_4	$(4, 5, 7, 8)_n$	$(1, 3, 6, 7)_n$	$(7, 8, 8, 9)_n$

Table 2 Decision matrix given by expert e_2

	A_1	A_2	A_3
C_1	$(3, 5, 6, 8)_n$	$(1, 2, 3, 4)_n$	$(2, 3, 4, 6)_n$
C_2	$(2, 3, 4, 5)_n$	$(3, 4, 5, 8)_n$	$(1, 3, 5, 8)_n$
C_3	$(2, 4, 5, 7)_n$	$(3, 4, 6, 7)_n$	$(1, 2, 4, 6)_n$
C_4	$(5, 6, 8, 9)_n$	$(1, 3, 6, 8)_n$	$(2, 3, 4, 6)_n$

Table 3 Decision matrix given by expert e_3

	A_1	A_2	A_3
C_1	$(5, 6, 7, 8)_n$	$(0, 1, 1, 2)_n$	$(2, 3, 4, 5)_n$
C_2	$(2, 4, 6, 7)_n$	$(4, 5, 6, 8)_n$	$(2, 4, 5, 7)_n$
C_3	$(5, 6, 8, 9)_n$	$(1, 4, 5, 8)_n$	$(3, 4, 6, 7)_n$
C_4	$(5, 7, 8, 9)_n$	$(3, 4, 5, 6)_n$	$(5, 6, 7, 8)_n$

Table 4 Normalized decision matrix given by expert e_1

	A_1	A_2	A_3
C_1	$(0.125, 0.333, 0.6, 1)_n$	$(0.5, 0.833, 1, 1)_n$	$(0.25, 0.667, 1, 1)_n$
C_2	$(0.333, 0.6, 1, 1)_n$	$(0.167, 0.6, 1, 1)_n$	$(0.333, 0.6, 1, 1)_n$
C_3	$(0.333, 0.571, 1, 1)_n$	$(0.444, 0.857, 1, 1)_n$	$(0.111, 0.429, 1, 1)_n$
C_4	$(0.444, 0.625, 0.875, 1)_n$	$(0.111, 0.375, 0.75, 1)_n$	$(0.778, 1, 1, 1)_n$

Table 5 Normalized decision matrix given by expert e_2

	A_1	A_2	A_3
C_1	$(0.375, 0.833, 1, 1)_n$	$(0.125, 0.333, 0.6, 1)_n$	$(0.25, 0.5, 0.8, 1)_n$
C_2	$(0.25, 0.6, 1, 1)_n$	$(0.375, 0.8, 1, 1)_n$	$(0.125, 0.6, 1, 1)_n$
C_3	$(0.286, 0.667, 1, 1)_n$	$(0.429, 0.667, 1, 1)_n$	$(0.143, 0.333, 1, 1)_n$
C_4	$(0.556, 0.75, 1, 1)_n$	$(0.111, 0.375, 1, 1)_n$	$(0.222, 0.375, 0.667, 1)_n$

Table 6 Normalized decision matrix given by expert e_3

	A_1	A_2	A_3
C_1	$(0.625, 0.857, 1, 1)_n$	$(0, 0.143, 0.167, 1)_n$	$(0.25, 0.429, 0.667, 1)_n$
C_2	$(0.25, 0.667, 1, 1)_n$	$(0.5, 0.833, 1, 1)_n$	$(0.25, 0.667, 1, 1)_n$
C_3	$(0.556, 0.75, 1, 1)_n$	$(0.111, 0.5, 0.833, 1)_n$	$(0.333, 0.5, 1, 1)_n$
C_4	$(0.556, 0.875, 1, 1)_n$	$(0.333, 0.5, 0.714, 1)_n$	$(0.556, 0.75, 1, 1)_n$

Table 7 Integrated decision matrix

	A_1	A_2	A_3
C_1	$(0.397, 0.691, 0.879, 1)_n$	$(0.190, 0.410, 0.514, 1)_n$	$(0.25, 0.523, 0.808, 1)_n$
C_2	$(0.275, 0.626, 1, 1)_n$	$(0.361, 0.753, 1, 1)_n$	$(0.237, 0.626, 1, 1)_n$
C_3	$(0.407, 0.671, 1, 1)_n$	$(0.309, 0.659, 0.934, 1)_n$	$(0.208, 0.428, 1, 1)_n$
C_4	$(0.522, 0.761, 0.962, 1)_n$	$(0.198, 0.424, 0.812, 1)_n$	$(0.522, 0.712, 0.898, 1)_n$

Step 2. Aggregate information by nonlinear operators of GFNs based on formula (5), formula (6) and formula (7), then get the integrated decision-making matrix as listed in Table 7.

Step 3. Calculate the comprehensive property values by weighted arithmetic average, we have

$V = ((0.413, 0.693, 0.964, 1)_n, (0.262, 0.558, 0.827, 1)_n, (0.316, 0.572, 0.931, 1)_n)^T$

Step 4. Calculate the mean values of GFNs based on formula (10), formula (11) and formula (12), we get

$$E(A_1) = \frac{0.2245n^2 + 0.9023n + 0.8294}{0.271n^2 + 1.129n + 1.174}$$

$$E(A_2) = \frac{0.1863n^2 + 0.8671n + 0.9314}{0.269n^2 + 1.276n + 1.476}$$

$$E(A_3) = \frac{0.2698n^2 + 1.02n + 0.9001}{0.359n^2 + 1.402n + 1.368}$$

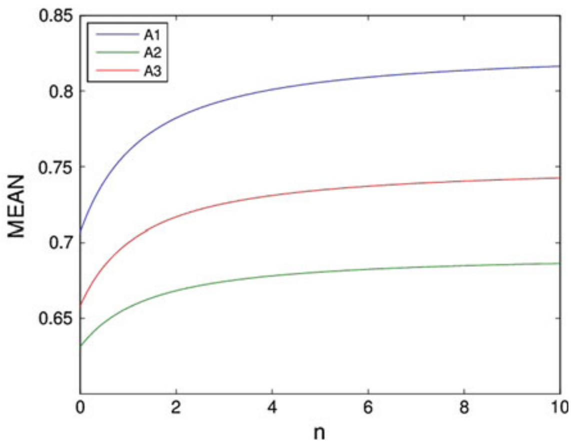
Step 5. Rank the alternatives based on the mean values of GFNs.

Based on the mean values, the ranking of the alternatives under different n values could be shown in Fig. 2.

According to the Fig. 2, for the decision makers of the car company, the ranking order is $A_1 \succ A_3 \succ A_2$.

At the same time, by comparing the sorted results under different n values, it could be shown that the final assessment results become more stable with the increase of n values. These results are suitable for the characteristics of the GFN, the robustness and adaptability of decision-making methods could be improved when the GFNs are considered in supplier selection. Moreover, the fuzziness of attribute is smaller and decision makers get more information with the increase of n values, therefore, the decision makers are easy to choose the best supplier.

Fig. 2 Ranking of the alternatives under different n



6 Conclusions

In the real decision problems, it is difficult to select the excellent supplier according to the different preference relations in fuzzy MCDM. In order to efficiently solve the complex decision making problems and find reasonable decision making methods, the fuzzy MCDM method with GFNs is proposed in the paper. Some nonlinear aggregation operators with generalized fuzzy numbers (GFNs) are also proposed. Furthermore, a ranking formula based on mean values of GFNs is adopted to rank alternatives. Finally, experimental results indicate that the proposed method offers a practical and effective way in fuzzy MCDM. It could be also shown that the final assessment results become more stable with the increase of n values. These results are suitable for the characteristics of the GFN. Furthermore, because the membership function of the GFN has different forms, the subjectivity of selecting the model parameters is reduced, and the robustness of the model is also improved.

Acknowledgments This research was supported in part by grants from the National Natural Science Foundation of China (#71222108, #71325001 and #71173028), Program for New Century Excellent Talents in University (#NCET-12-0086), the Research Fund for the Doctoral Program of Higher Education (#20120185110031) the Fundamental Research Funds for the Central Universities (ZYGX2015KYQD079).

References

1. Aissaoui, N., Haouari, M., Hassini, E.: Supplier selection and order lot sizing modeling: a review. Comput. Oper. Res. **34**(12), 3516–3540 (2007)
2. Kou, G., Ergu, D., Shang, J.: Enhancing data consistency in decision matrix: adapting hadamard model to mitigate judgment contradiction. Eur. J. Oper. Res. **236**(1), 261–271 (2014)
3. Kou, G., Ergu, D., Shi, Y.: An integrated expert system for fast disaster assessment. Comput. Oper. Res. **42**, 95–107 (2014)
4. Li, G., Kou, G., Peng, Y.: Dynamic fuzzy multiple criteria decision making for performance evaluation. Technol. Econ. Dev. Econ. **21**(5), 705–719 (2015)
5. Jadidi, O., Zolfaghari, S., Cavalieri, S.: A new normalized goal programming model for multi-objective problems: a case of supplier selection and order allocation. Int. J. Prod. Econ. **148**, 158–165 (2014)
6. Ho, W., Xu, X., Dey, P.K.: Multi-criteria decision making approaches for supplier evaluation and selection: a literature review. Eur. J. Oper. Res. **202**(1), 16–24 (2010)
7. Zadeh, L.A.: Fuzzy sets. Inform. Control **8**(3), 338–353 (1965)
8. Chen, C.-T., Lin, C.-T., Huang, S.-F.: A fuzzy approach for supplier evaluation and selection in supply chain management. Int. J. Prod. Econ. **102**(2), 289–301 (2006)
9. Junior, F.R.L., Osiro, L., Carpinetti, L.C.R.: A comparison between fuzzy ahp and fuzzy topsis methods to supplier selection. Appl. Soft Comput. **21**, 194–209 (2014)
10. Senvar, O., Tuzkaya, G., Kahraman, C.: Multi criteria supplier selection using fuzzy promethee method. In: Supply Chain Management Under Fuzziness, pp. 21–34. Springer (2014)
11. Li, G., Kou, G., Lin, C., Xu, L., Liao, Y.: Multi-attribute decision making with generalized fuzzy numbers. J. Oper. Res. Soc. **66**(11), 1793–1803 (2015)
12. Wu, D.D., Zhang, Y., Wu, D., Olson, D.L.: Fuzzy multi-objective programming for supplier selection and risk modeling: a possibility approach. Eur. J. Oper. Res. **200**(3), 774–787 (2010)

13. Lin, R.-H.: An integrated model for supplier selection under a fuzzy situation. Int. J. Prod. Econ. **138**(1), 55–61 (2012)
14. Nazari-Shirkouhi, S., Shakouri, H., Javadi, B., Keramati, A.: Supplier selection and order allocation problem using a two-phase fuzzy multi-objective linear programming. Appl. Math. Model. **37**(22), 9308–9323 (2013)
15. Yager, R.R., Kacprzyk, J.: The ordered weighted averaging operators: theory and applications, Springer Science and Business Media (2012)
16. Chiclana, F., Herrera, F., Herrera-Viedma, E.: Integrating multiplicative preference relations in a multipurpose decision-making model based on fuzzy preference relations. Fuzzy Sets Syst. **122**(2), 277–291 (2001)
17. Xu, Z.-S., Da, Q.-L.: The uncertain owa operator. Int. J. Intell. Syst. **17**(6), 569–575 (2002)
18. Zhou, L.-G., Chen, H.-Y.: Generalized ordered weighted logarithm aggregation operators and their applications to group decision making. Int. J. Intell. Syst. **25**(7), 683–707 (2010)
19. Yager, R.R.: The power average operator. IEEE Trans. Syst. Man Cybern. Part A: Syst. Humans **31**(6), 724–731 (2001)
20. Xu, Z., Yager, R.R.: Power-geometric operators and their use in group decision making. IEEE Trans. Fuzzy Syst. **18**(1), 94–105 (2010)
21. Ergu, D., Kou, G., Shi, Y., Shi, Y.: Analytic network process in risk assessment and decision analysis. Comput. Oper. Res. **42**, 58–74 (2014)
22. Asady, B.: The revised method of ranking lr fuzzy number based on deviation degree. Expert Syst. Appl. **37**(7), 5056–5060 (2010)
23. Matarazzo, B., Munda, G.: New approaches for the comparison of l-r fuzzy numbers: a theoretical and operational analysis. Fuzzy Sets Syst. **118**(3), 407–418 (2001)

Possibility-Based Outranking Comparison for PROMETHEE II with Uncertain Linguistic Fuzzy Variables

Binquan Liao, Lidong Wang and Xiaodong Liu

Abstract Linguistic variable is an effective tool to represent the complexity and uncertainty of alternatives information, which closes to the fuzziness of man's thinking. In this paper, PROMETHEE II-based ranking method is studied under the environment of uncertain linguistic fuzzy variables. The preference measure functions are refined by introducing the overall dominance possibility assignment methods among uncertain linguistic fuzzy variables, by which the exiting flow, entering flow and net flow are obtained respectively. Additionally, the rank of alternative can be achieved easily by comparing net flow. The convenience and applicability of the proposed methods are illustrated with a practical example about the evaluation of socio-economic systems.

Keywords Uncertain linguistic fuzzy variables · The overall dominance possibility · PROMETHEE II method

1 Introduction

Multiple attribute decision making (MADM) problems exist widely in real-life decision making situation [1–5]. For traditional MADM problems, decision makers assess each alternative under each attribute expressing with numerical values. However, numerical values are not enough to reflect the human cognitive mechanism. Uncertain knowledge exist broadly in daily life and the process of decision making, it is suitable for decision-makers to give the evaluation information by virtue of linguistic terms as opposed to precise values [6]. For instance, a decision maker may

B. Liao · L. Wang (✉) · X. Liu
Department of Mathematics, Dalian Maritime University, Dalian 116026, People's Republic of China
e-mail: ldwang@hotmail.com

L. Wang
Artificial Intelligence Key Laboratory of Sichuan Province, Sichuan University of Science and Engineering, Zigong 643000, People's Republic of China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_8

evaluate the financial environment of a city with the fuzzy linguistic term "good" instead of numerical values 85. MADM problem under uncertain linguistic information has been a hot research topic and attracted more and more researcher to study it during the last several years.

Due to the complexity of certain decision-making problems, the evaluation information provided by decision-makers may be exists between two linguistic phrases [6]. For instance, experts provide his/her preferences over alternatives with range from a linguistic label to another. In order to deal with this case, Xu defined the uncertain linguistic variables and introduced some of their operational laws [7, 8], in which experts can provide quantitative information expressed by a pair of linguistic variables. In group linguistic decision analysis, except the choice and operations of the linguistic term set, how to establish an appropriate aggregation operator is a necessary step, by which one can obtain comprehensive evaluation result derived from multiple experts [9]. Xu developed the uncertain linguistic ordered weighted averaging (ULOWA) operator and uncertain linguistic hybrid aggregation (ULHA) operator [10]. Xu introduced a distance measures between two uncertain linguistic fuzzy sets, and solved the problem under different types of linguistic information [11]. Yue and Fan provided the property of the dominance degree on pairwise comparisons of uncertain linguistic variables [12]. Peng developed an uncertain pure linguistic hybrid harmonic averaging operator and a generalized interval aggregation operator, and applied it in an investment company that invest a sum of money in the best option [13]. Meng and Tan discussed the Choquet aggregation operators ranking decision under uncertain linguistic environment [14].

The likelihood-based assignment was introduced in the MADM problem, in which the evaluative of the alternatives and the importance weights of the attribute are expressed in interval fuzzy numbers [15]. Likelihood-based assignment denotes the evaluative rating of alternative is not smaller than the evaluative rating of another alternative with respect to criterion, which can avoided the loss and distortion of information, so the feasibility and the applicability of this method can solve the MADM problem. Chen established an interval type-2 fuzzy PROMETHEE method using lower and upper likelihoods-based outranking index [15]. Wang and Chen employed likelihoods-based outranking methods to interval-valued intuitionistic fuzzy problem [16].

Inspired by work above mentioned, this paper introduces overall dominance possibility assignment methods to a MADM problem with uncertain linguistic information. The rest of this paper is organized as follows. Section 2. describes the concepts of uncertain linguistic variables. In Sect. 3, an improved PROMETHEE II method is described under uncertain linguistic fuzzy sets by introducing overall dominance possibility assignment method. In Sect. 4, an example is provided for showing the feasibility and applicability of the proposed methodology. Finally, the conclusions are draw in Sect. 5.

2 Preliminaries

In this section, some related concepts and operators of the uncertain linguistic variables are recalled.

It is necessary to choose an appropriate linguistic labels set in advance. Let $S = \{s| = 0, 1, \ldots, L-1\}$, where s represents a linguistic variable, L is odd number [10]. In general, L can takes 3, 5, 7, 9 and so on. The S has the following characteristics [10].

(1) If $i > j$, then $s_i > s_j$ (i.e., s_i is better than s_j),
(2) There exists a negative operator $neg(s_i) = s_{L-1-i}$,
(3) If $s_i \geq s_j$ (i.e., s_i is not worse than s_j), then $\max(s_i, s_j) = s_i$,
(4) If $s_i \leq s_j$ (i.e., s_i is not better than s_j), then $\min(s_i, s_j) = s_i$.

As is pointed by Xu [10], in order to minimize the linguistic information loss during the operational process, the original discrete linguistic assessment set $S = (s_0, s_1, \ldots, s_{L-1})$ is extended to continuous linguistic assessment set $\bar{s} = \{s_\alpha | \alpha \in [0, q]\}$, where q is a sufficiently large number. If $s_\alpha \in S$, then s_α is called a original linguistic term, or else s_α is called an extended linguistic term [10].

Due to the experts' uncertain knowledge about the preference degrees of one alternative over another, linguistic preference information provided by the experts presented interval linguistic values is better than that of precise linguistic variables. In the following paragraphs, some related concepts are introduced. Suppose $\overline{S}$ is a set of all uncertain linguistic variables, $\tilde{s} = [s_a, s_b] \in \overline{S}$ where $a \leq b$, s_a, s_b are lower and upper limit of $\tilde{s}$ respectively, then $\tilde{s}$ is called an uncertain linguistic variable.

Definition 1 ([11]) Suppose $\overline{S}$ is a set of all uncertain linguistic variables, and for any $\tilde{s}_1 = [s_{a1}, s_{b1}]$, $\tilde{s}_2 = [s_{a2}, s_{b2}] \in \overline{S}$, then the operational rules are shown as follows:

(1) $\tilde{s}_1 \oplus \tilde{s}_2 = [s_{a1}, s_{b1}] \oplus [s_{a2}, s_{b2}] = [s_{a1+a2}, s_{b1+b2}]$,
(2) $\tilde{s}_1 \otimes \tilde{s}_2 = [s_{a1}, s_{b1}] \otimes [s_{a2}, s_{b2}] = [s_{a1\times a2}, s_{b1\times b2}]$,
(3) $\tilde{s}_1 / \tilde{s}_2 = [s_{a1}, s_{b1}] / [s_{a2}, s_{b2}] = [s_{a1/b2}, s_{b1/a2}]$, if $a_2 \neq 0$, $b_2 \neq 0$,
(4) $\lambda\tilde{s}_1 = [s_{\lambda * a1}, s_{\lambda * b1}], \quad \lambda \geq 0$,
(5) $\lambda(\tilde{s}_1 \oplus \tilde{s}_1) = \lambda\tilde{s}_1 \oplus \lambda\tilde{s}_2, \quad \lambda \geq 0$,
(6) $(\lambda_1 + \lambda_2)\tilde{s}_1 = \lambda\tilde{s}_1 \oplus \lambda\tilde{s}_1, \quad \lambda_1, \lambda_2 \geq 0$.

Definition 2 ([17]) Let $\tilde{s}_1 = [s_{a_1}, s_{b_1}]$ and $\tilde{s}_2 = [s_{a_2}, s_{b_2}]$ be two uncertain linguistic variables, and $l_{\tilde{s}_1} = b_1 - a_1$ and $l_{\tilde{s}_2} = b_2 - a_2$, then the degree of possibility of $\tilde{s}_1 \geq \tilde{s}_2$ is defined as:

$$p(\tilde{s}_1 \geq \tilde{s}_2) = \frac{\max\{0, l_{\tilde{s}_1} + l_{\tilde{s}_2} - \max(b_2 - a_1, 0)\}}{l_{\tilde{s}_1} + l_{\tilde{s}_2}}. \tag{1}$$

From Definition 2, one can get the following results:

(1) $0 \leq p(\tilde{s}_1 \geq \tilde{s}_2) \leq 1, 0 \leq p(\tilde{s}_2 \geq \tilde{s}_1) \leq 1$,
(2) $p(\tilde{s}_1 \geq \tilde{s}_2) + p(\tilde{s}_2 \geq \tilde{s}_1) = 1$. Especially, $p(\tilde{s}_1 \geq \tilde{s}_1) = p(\tilde{s}_2 \geq \tilde{s}_2) = 1/2$.

For the two interval fuzzy numbers, the degree of possibility of them is defined as follows:

Definition 3 ([18]) Let $a = [a^-, a^+]$ and $b = [b^-, b^+]$ be any two interval fuzzy numbers, then the degree of possibility of $a \geq b$ is defined as:

$$\widetilde{p}(a \geq b) = \frac{\max(0, a^+ - b^-) - \max(0, a^- - b^+)}{a^+ - a^- + b^+ - b^-} \tag{2}$$

3 Possibility-Based Comparison for PROMETHEE II Model

This section establishes a PROMETHEE II method involving the interval weight information by introducing the degree of overall dominance possibility index of the uncertain linguistic preference relations.

3.1 Uncertain Linguistic Decision System

Considering a multiple attribute decision making with uncertain linguistic information: let $X = \{x_1, x_2, \ldots, x_n\}$ be a discrete set of alternatives, and $U = \{u_1, u_2, \ldots, u_m\}$ be the set of attributes, $W = \{w_1, w_2, \ldots, w_m\}$ be the interval weight vector of attribute, where $w_j = [w_j^-, w_j^+], w_j^-, w_j^+ > 0, w_j^- < w_j^+$. Suppose that $\tilde{A} = (A_{ij})_{n \times m}$ is decision matrix, where $A_{ij} \in \bar{S}(A_{ij} = [A_{ij}^-, A_{ij}^+], A_{ij}^- \in S, A_{ij}^+ \in S)$, which takes the form of uncertain linguistic variable given by the decision maker for alternative $x_i \in X$ with respect to the attribute $u_j \in U$.

3.2 The Overall Dominance Possibility Assignment Methods

The index $p(A_{ij} \geq A_{i'j})$ only reflects the degree of dominance possibility between the two alternatives, and it cannot reflect the degree of overall dominance possibility between one alternative and other alternatives under an attribute. Inspired by the Chen's work [15], the degree of overall dominance possibility index $OP(A_{ij})$ is defined under the uncertain linguistic environment. It can reflect the overall degree of possibility between one alternative and other alternatives with respect to each attribute.

Definition 4 Assume that A_{ij} and $A_{i'j}$ are two ULF evaluative information of the alternatives $x_i, x_{i'} \in X$ with respect to the attribute $u_j \in U$. The degree of overall dominance possibility index $OP(A_{ij})$ of $A_{ij} \in S$ is defined as follows:

$$OP(A_{ij}) = \frac{1}{n-1} \sum_{\substack{i'=1 \\ i' \neq i}}^{n} p\left(A_{ij} \geq A_{i'j}\right) \tag{3}$$

where $p(A_{ij} \geq A_{i'j})$ defined by Eq. (1).

3.3 Dominance Possibility Preference Functions for PROMETHEE II Model

The PROMETHEE II is a powerful and effective method dealed with MADM problem. There are six shapes of preference functions in the original PROMETHEE II methods, which proposed by Brans [19], Brans and Vincke [20], and Brans et al. [21]. Classical preference functions derived from the differences between alternatives for a certain attribute, which bring some difficulties of the comparison between the two of the evaluations rating under fuzzy linguistic environment. The likelihood-based outing index $LI(A)_{ij}$ is introduced to define new preference functions by Chen [15], which shows that the comparison between two of the evaluations rating is obtained by introducing the relative differences between alternatives for a certain attribute.

3.3.1 Overall Dominance Possibility Preference Function

In this part, the overall dominance possibility preference function is employed to replace the traditional preference function under ULF evaluation rating and interval weight. This preference function integrates the difference between the two degree of overall dominance possibility index. Then we obtain the alternative rank by using PROMETHEE II methods.

Inspired by the Chen's work [15], the overall dominance possibility Gaussian preference function $P^{VI}(A_{ij} \succ A_{i'j})$ is described specifically as follows.

Definition 5 Let $OP(A_{ij})$ and $OP(A_{i'j})$ be two overall dominance possibility indices of the ULF evaluative ratings A_{ij} and $A_{i'j}$ for the alternatives $x_i, x_{i'} \in X$ with respect to the criterion $u_j \in U$, respectively. The overall dominance possibility Gaussian preference function $P^{VI}(A_{ij} \succ A_{i'j})$ of A_{ij} with respect to $A_{i'j}$ is defined as follows:

The overall dominance possibility Gaussian preference function:

$$P^{VI}(A_{ij} \succ A_{i'j}) = \begin{cases} 1 - e^{-(OP(A_{ij}) - OP(A_{i'j}))^2/2\sigma^2} & \text{if } \mathrm{OP(A_{ij}) - OP(A_{i'j}) > 0} \\ 0 & \text{if } \mathrm{OP(A_{ij}) - OP(A_{i'j}) \leq 0.} \end{cases} \tag{4}$$

3.3.2 Comprehensive Preference Measures

In this section, a comprehensive preference measure $\Pi_{ii'}$ is constructed for each pair of alternatives x_i and $x_{i'}$ based on the overall dominance possibility preference function and the interval weight of attribute.

Definition 6 For each pair of alternatives x_i and $x_{i'} \in X(i, i' = 1, 2, \ldots, n$ and $i \neq i')$, the comprehensive preference measure $\Pi_{ii'}$ of x_i with respect to $x_{i'}$ is defined as follows:

$$\Pi_{ii'} = \left(\oplus_{j=1}^{m} P^{VI}\left(A_{ij} \succ A_{i'j}\right) \cdot w_j\right) / \left(\oplus_{j=1}^{m} w_j\right), \tag{5}$$

where

$$\begin{aligned}
\Pi_{ii'} &= [\Pi_{ii'}^{-}, \Pi_{ii'}^{+}] \\
\Pi_{ii'}^{-} &= \textstyle\sum_{j=1}^{m} P^{VI}(A_{ij} \succ A_{i'j}) \cdot w_j^{-} / \textstyle\sum_{j=1}^{m} w_j^{+} \\
\Pi_{ii'}^{+} &= \textstyle\sum_{j=1}^{m} P^{VI}(A_{ij} \succ A_{i'j}) \cdot w_j^{+} / \textstyle\sum_{j=1}^{m} w_j^{-}.
\end{aligned}$$

Notice that $x_i, x_{i'} \in X$ and $i \neq i'$, and $\Pi_{ii'}$ is an interval number.

3.3.3 ULF Flows

In this part, in order to determine the outranking relationships, the comprehensive preference measures are applied to define the ULF exiting flow, the ULF entering flow, and the ULF net flow.

Definition 7 For any $x_i, x_{i'} \in X$, the ULF exiting flow $\boldsymbol{\Phi}^{L}$, entering flow $\boldsymbol{\Phi}^{E}$, and net flow $\boldsymbol{\Phi}^{N}$ for each alternative $x_i \in X$ are defined as follows:

$$\begin{aligned}
\boldsymbol{\Phi}_i^{L} &= \oplus_{i'=1, i' \neq i}^{n} \boldsymbol{\Pi}_{ii'} \\
\boldsymbol{\Phi}_i^{E} &= \oplus_{i'=1, i' \neq i}^{n} \boldsymbol{\Pi}_{i'i} \\
\boldsymbol{\Phi}_i^{N} &= \boldsymbol{\Phi}_i^{L} \Theta \boldsymbol{\Phi}_i^{E} = [\boldsymbol{\Phi}_i^{N^-}, \boldsymbol{\Phi}_i^{N^+}],
\end{aligned} \tag{6}$$

where $\boldsymbol{\Phi}_i^{N^-} = \boldsymbol{\Phi}_i^{L^-} - \boldsymbol{\Phi}_i^{E^+}$, $\boldsymbol{\Phi}_i^{N^+} = \boldsymbol{\Phi}_i^{L^+} - \boldsymbol{\Phi}_i^{E^-}$, $\prod_{ii'}$ is the comprehensive preference measure of x_i with respect to $x_{i'}$, where $x_i, x_{i'} \in X$.

Definition 8 For any $x_i \in X$, the comprehensive outranking index $\boldsymbol{\Phi I}^{L}(x_i)$, the comprehensive outranked index $\boldsymbol{\Phi I}^{E}(x_i)$, and the comprehensive dominance index $\boldsymbol{\Phi I}^{N}(x_i)$ of x_i are defined, respectively, as follows:

$$\begin{aligned}
\boldsymbol{\Phi I}^L(x_i) &= \frac{1}{(n-1)} \sum_{\substack{i'=1 \\ i' \neq i}}^{n} \widetilde{p}(\boldsymbol{\Phi}_i^L \geq \boldsymbol{\Phi}_{i'}^L) \\
\boldsymbol{\Phi I}^E(x_i) &= \frac{1}{(n-1)} \sum_{\substack{i'=1 \\ i' \neq i}}^{n} \widetilde{p}(\boldsymbol{\Phi}_i^E \geq \boldsymbol{\Phi}_{i'}^E) \\
\boldsymbol{\Phi I}^N(x_i) &= \frac{1}{(n-1)} \sum_{\substack{i'=1 \\ i' \neq i}}^{n} \widetilde{p}(\boldsymbol{\Phi}_i^N \geq \boldsymbol{\Phi}_{i'}^N).
\end{aligned} \tag{7}$$

where $\boldsymbol{\Phi}_i^L$, $\boldsymbol{\Phi}_i^E$ and $\boldsymbol{\Phi}_i^E$ denote the ULF exiting flow, the ULF entering flow, and the ULF net flow, respectively. $\widetilde{p}(\boldsymbol{\Phi}_i^L \geq \boldsymbol{\Phi}_{i'}^L)$ represents the degree of possibility of two interval fuzzy numbers defined as Eq. (2).

For any $x_i \in X$, the following properties hold:

$$0 \leq \boldsymbol{\Phi I}^L(x_i) \leq 1, 0 \leq \boldsymbol{\Phi I}^E(x_i) \leq 1, 0 \leq \boldsymbol{\Phi I}^N(x_i) \leq 1.$$

3.3.4 PROMETHEE II-Based Ranking

The comprehensive dominance index $\boldsymbol{\Phi I}^N(x_i)$ is employed to rank the alternatives in a total preorder, the larger $\boldsymbol{\Phi I}^N(x_i)$, the more the alternative x_i dominates the other alternatives in X [22]. The ULF PROMETHEE II complete relationships with respect to an MADM problem.

3.4 Steps of Proposed Model

According to the above mentioned analysis, detailed steps are given to solve the ULF MADM problem.

Step 1: Establish the decision matrix of $\tilde{A} = (A_{ij})_{n \times m}$ with the uncertain linguistic evaluative rating for each alternative $x_i \in X$ with respect to the attribute $u_j \in U$.

Step 2: Determine the interval weight w_j for the attribute $u_j \in U$.

Step 3: Calculate the degree of possibility $p(A_{ij} \geq A_{i'j})$ with respect to $u_j \in U$ for each alternative pair $(x_i, x_{i'})$.

Step 4: Calculate the overall dominance possibility index $OP(A_{ij})$ of the uncertain linguistic evaluative ratings A_{ij}.

Step 5: Specify the corresponding parameters σ, then calculate the overall dominance possibility Gaussian preference function $P^{VI}(A_{ij} \succ A_{i'j})$.

Step 6: Compute the comprehensive preference measure $\boldsymbol{\Pi}_{ii'}$ for each of $(x_i, x_{i'})$ where $x_i \in X, x_{i'} \in X$ and $i \neq i'$.

Step 7: Calculate the uncertain fuzzy exiting flow $\boldsymbol{\Phi}_i^L$, the uncertain fuzzy entering flow $\boldsymbol{\Phi}_i^E$, and the uncertain fuzzy net flow $\boldsymbol{\Phi}_i^N$ for each alternative $x_i \in X$.

Step 8: Calculate the comprehensive dominance index $\boldsymbol{\Phi I}^N(x_i)$ for each alternative $x_i \in X$, by which the complete ranking orders can be obtained for all alternatives in X.

4 Illustrative Applications

In this section, an example for evaluation of socio-economic systems in reference [7] is employed to verify the feasibility and applicability of the proposed method. The evaluation of socio-economic systems including five candidate city is denoted by $X = \{x_1, x_2, x_3, x_4, x_5\}$, described by attribute set $U = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9\}$ with ULF language description $S = \{s_i | i \in [0, 10]\}$, where political environment u_1, economic environment u_2, financial environment u_3, administrative environment u_4, market environment u_5, technical condition u_6, material basis u_7, legal environment u_8, and natural environment u_9 [7]. Uncertain information of experts decision are listed in Tables 1 and 2.

This paper combines the degree of the overall dominance possibility and Gaussian preference function to reflect the difference of that alternatives, in which the parameter σ takes 0.3.

Step 1. The decision matrix of $\tilde{A} = (A_{ij})_{n \times m}$ showed in Table 1.

Step 2. Interval weight information is determine as Table 2.

Step 3. Calculate the $p(A_{ij} \geq A_{i'j})$ for each pair of alternatives $x_i, x_{i'}$ in regard to each $u_j \in U$. For example, measure the ULF binary relationship $A_{32} \geq A_{22}$

$$p(A_{32} \geq A_{22}) = \frac{\max\{(8-6) + (7-5) - \max(7-6, 0), 0\}}{(8-6) + (7-5)} = 0.75$$

Step 4. Calculate the overall dominance possibility index $OP(A_{ij})$ for $x_i \in X$ and $u_j \in U$, and the results are showed in Table 3.

Step 5. According to the overall dominance possibility preference function mentioned above, one can obtain the preference measure $P^{VI}(A_{ij} \succ A_{i'j})$ for any $i, i' \in \{1, 2, 3, 4, 5\}, j \in \{1, 2, \ldots, 9\}$. For instances,

Table 1 Uncertain linguistic decision matrix [7]

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9
x_1	$[s_7, s_8]$	$[s_8, s_9]$	$[s_8, s_9]$	$[s_8, s_9]$	$[s_6, s_7]$	$[s_8, s_9]$	$[s_5, s_8]$	$[s_8, s_9]$	$[s_8, s_9]$
x_2	$[s_6, s_9]$	$[s_5, s_7]$	$[s_8, s_9]$	$[s_6, s_7]$	$[s_8, s_{10}]$	$[s_8, s_9]$	$[s_8, s_9]$	$[s_9, s_{10}]$	$[s_7, s_8]$
x_3	$[s_6, s_7]$	$[s_6, s_8]$	$[s_7, s_8]$	$[s_6, s_7]$	$[s_7, s_9]$	$[s_6, s_8]$	$[s_8, s_9]$	$[s_8, s_9]$	$[s_7, s_8]$
x_4	$[s_6, s_7]$	$[s_8, s_9]$	$[s_7, s_8]$	$[s_8, s_9]$	$[s_6, s_8]$	$[s_7, s_8]$	$[s_9, s_{10}]$	$[s_7, s_9]$	$[s_6, s_8]$
x_5	$[s_6, s_7]$	$[s_6, s_9]$	$[s_7, s_{10}]$	$[s_7, s_8]$	$[s_7, s_{10}]$	$[s_7, s_8]$	$[s_6, s_7]$	$[s_6, s_8]$	$[s_8, s_9]$

Table 2 Weight information [7]

u_1	u_2	u_3	u_4	u_5
[0.08, 0.10]	[0.05, 0.09]	[0.10, 0.12]	[0.08, 0.11]	[0.10, 0.13]
u_6	u_7	u_8	u_9	
[0.12, 0.14]	[0.14, 0.16]	[0.09, 0.11]	[0.11, 0.15]	

Table 3 The degree of overall dominance possibility index

	$OP(A_{i1})$	$OP(A_{i2})$	$OP(A_{i3})$	$OP(A_{i4})$	$OP(A_{i5})$	$OP(A_{i6})$	$OP(A_{i7})$	$OP(A_{i8})$	$OP(A_{i9})$
x_1	0.88	0.81	0.75	0.88	0.08	0.88	0.13	0.54	0.88
x_2	0.69	0.11	0.75	0.13	0.84	0.88	0.63	1	0.29
x_3	0.31	0.29	0.19	0.13	0.6	0.17	0.63	0.54	0.29
x_4	0.31	0.81	0.19	0.88	0.28	0.29	1	0.35	0.17
x_5	0.31	0.48	0.63	0.5	0.7	0.29	0.13	0.06	0.88

Table 4 Comprehensive preference measures $\Pi_{ii'}$

Π_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
$i = 1$	______	[0.21, 0.39]	[0.43, 0.71]	[0.34, 0.53]	[0.28, 0.45]
$i = 2$	[0.24, 0.37]	______	[0.3, 0.46]	[0.36, 0.56]	[0.32, 0.49]
$i = 3$	[0.16, 0.25]	[0.01, 0.02]	______	[0.06, 0.10]	[0.15, 0.23]
$i = 4$	[0.14, 0.21]	[0.18, 0.32]	[0.18, 0.31]	______	[0.21, 0.34]
$i = 5$	[0.08, 0.13]	[0.15, 0.27]	[0.20, 0.35]	[0.21,0.35]	______

Table 5 Exiting flow Φ_i^L, entering flow Φ_i^E, net flow Φ_i^N

	Φ_i^L	Φ_i^E	Φ_i^N
x_1	[1.26, 2.08]	[0.62, 0.96]	[0.30, 1.46]
x_2	[1.22, 1.88]	[0.55, 1]	[0.22, 1.33]
x_3	[0.38, 0.60]	[1.11, 1.83]	[–1.45, –0.51]
x_4	[0.71, 1.18]	[0.97, 1.54]	[–0.83, 0.21]
x_5	[0.64, 1.10]	[0.96, 1.51]	[–0.87, 0.14]

$$
\begin{aligned}
P^{VI}(A_{12} \succ A_{22}) &= 1 - e^{-(OP(A_{12})-OP(A_{22}))^2/2\sigma^2} \\
&= 1 - e^{-(0.81-0.11)^2/2\cdot 0.3^2} \\
&= 0.93.
\end{aligned}
$$

Step 6. Combining preference measure and interval weights, one can get the comprehensive preference measures $\Pi_{ii'}$ for $x_i, x_{i'} \in X$ and $u_j \in U$, and the results are showed in Table 4.

Step 7. Calculate the ULF exiting flow Φ_i^L, the ULF entering flow Φ_i^E and the ULF net flow Φ_i^N for each $x_i \in X$. The results are showed in Table 5.

The comprehensive dominance index is obtained in Table 6. By the comprehensive dominance index, one can obtain the complete ranking orders of all alternative as follows:

$$x_1 \succ x_2 \succ x_4 \succ x_5 \succ x_3$$

So, the alternative x_1 is the best scheme selection.

Table 6 Exiting flow Φ_i^L, entering flow Φ_i^E, net flow Φ_i^N

	$\Phi I^L(x_i)$	$\Phi I^E(x_i)$	$\Phi I^N(x_i)$
x_1	0.90	0.13	0.89
x_2	0.85	0.14	0.86
x_3	0	0.84	0.09
x_4	0.40	0.71	0.34
x_5	0.35	0.69	0.32

5 Conclusion

In this paper, PROMETHEE II based on ranking method is established under the uncertain linguistic information and interval weight. By introducing overall dominance possibility assignment methods among uncertain linguistic fuzzy variables, exiting flow, entering flow and net flow are obtained with the form of interval number. The rank of alternatives is achieved by employing comprehensive dominance index. A practical example is given to demonstrate its practicality and effectiveness for the evaluation of socio-economic systems.

Acknowledgments This paper is supported by the Natural Science Foundation of China (Nos. 61203283, 61305100), Liaoning Provincial Natural Science Foundation of China (No. 2014025004), Scientific Research Project of Liaoning Provincial Education Department (No. L2015072) and the Fundamental Research Funds for the Central Universities (Nos.3132014036, 3132014324).

References

1. Merigo, J.M.: The uncertain probabilistic weighted average and its application in the theory of expertons. Afr. J. Bus. Manage. **5**, 6092–6102 (2011)
2. Merig, J.M., Casanovas, M., Martnez, L.: Linguistic aggregation operators for linguistic decision making based on the Dempster-Shafer theory of evidence. Int. J. Uncertainty, Fuzziness Knowl.-Based Syst. **18**(3), 287–304 (2010)
3. Wei, G.W., Zhao, X.F., Lin, R., Wang, H.J.: Generalized triangular fuzzy correlated averaging operator and their application to multiple attribute decision making. Appl. Math. Model. **36**, 2975–2982 (2012)
4. Wei, G.W.: GRA-based linear-programming methodology for multiple attribute group decision making with 2-tuple linguistic assessment information. Inform. Int. Interdisc. J. **14**(4), 1105–1110 (2011)
5. Yue, Z.L.: A method for group decision-making based on determining weights of decision makers using TOPSIS. Appl. Math. Model. **35**, 1926–1936 (2011)
6. Wang, J.Q., Peng, J.J., Zhang, H.Y., Liu, T., Chen, X.H.: An uncertain linguistic multi-criteria group decision-making method based on a cloud model. Group Decis. Negot. **24**, 171–192 (2015)
7. Xu, Z.S.: Uncertain Multi-Attribute Decision Making: Methods and Applications. Springer (2015)

8. Xu, Z.S.: A method based on linguistic aggregation operators for group decision making with linguistic preference relations. Inform. Sci. **166**(1), 19–30 (2004)
9. Huynh, V.N., Nakamori, Y.: Multi-expert decision-making with linguistic information: a probabilistic-based model. In: Proceedings of the 38th Hawaii International Conference on System Sciences, pp. 1–9 (2005)
10. Xu, Z.S.: Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment. Inf. Sci. **168**, 171–184 (2004)
11. Xu, Z.S.: Induced uncertain linguistic OWA operators applied to group decision making. Inform. Fusion **7**(2), 231–238 (2006)
12. Yue, Q., Fan, Z.P.: Method for solving multiple attribute group decision-making problems with multigranularity uncertain linguistic assessment information. Control Decis. **25**(7), 1059–1062 (2010)
13. Peng, B., Ye, C., Zeng, S.Z.: Uncertain pure linguistic hybrid harmonic averaging operator and generalized interval aggregation operator b ased approach to group decision making. Knowl.-Based Syst. **36**, 175–181 (2012)
14. Meng, F., Tan, C.Q., Zhang, Q.: An approach to multi-attribute group decision making under uncertain linguistic environment based on the Choquet aggregation operators. J. Intell. Fuzzy Syst. **26**, 769–780 (2014)
15. Chen, T.Y.: An interval type-2 fuzzy PROMETHEE method using a likelihood-based outranking comparison approach. Inform. Fusion **25**, 105–120 (2015)
16. Wang, J.C., Chen, T.Y.: Likelihood-based assignment methods for multiple criteria decision analysis based on interval-valued intuitionistic fuzzy sets. Fuzzy Optim. Decis. Making (2015). doi:10.1007/s10700-015-9208-6
17. Xu, Z.S.: An approach based on the uncertain LOWG and induced uncertain LOWG operators to group decision making with uncertain multiplicative linguistic preference relations. Decis. Support Syst. **41**, 488–499 (2006)
18. Wang, Y.M., Yang, J.B., Xu, D.L.: A two-stage logarithmic goal programming method for generating weights from interval comparison matrices. Fuzzy Sets Syst. **152**, 475–498 (2005)
19. Brans, J.P.: L'ingéniérie de la décision: Élaboration d'instruments d'aide á la décision. La méthode PROMETHEE. In: R. Nadeau, M. Landry (Eds.), L'aide á la décision: Nature, Instruments et Perspectives d'Avenir, Presses de 1'Unoversoté Laval, Québec, pp. 183–213 (1982)
20. Brans, J.P., Vincke, P.H.: A preference ranking organization method (the PROMETHEE method for multiple criteria decision making). Manage. Sci. **31**(6), 647–656 (1985)
21. Brans, J.P., Vincke, P.H., Mareschal, B.: How to select and how to rank projects: the PROMETHEE method. Eur. J. Oper. Res. **24**(2), 228–238 (1986)
22. Betrie, G.D., Sadiq, R., Morin, K.A., Tesfamariam, S.: Selection of remedial alternatives for mine sites: a multicriteria decision analysis approach. J. Environ. Manage. **119**, 36–46 (2013)

Hesitant Fuzzy Correlation Measures Considering the Credibility

Jian-hui Yang and Chuan-yang Ruan

Abstract Presenting corresponding credibility on every hesitant fuzzy evaluation value, According to the hesitant fuzzy comprehensive evaluation problem of unknown attribute weights information, we define several hesitant fuzzy correlation formulas considering credibility. Then we construct several optimal attribute weighs model based on correlation for attribute weights.

Keywords Hesitant fuzzy · Correlation measures · Credibility

1 Introduction

Torra [1, 2] introduced the hesitant fuzzy set, which allows an element belonging to a set's membership grade can be different values. Currently, multi-attribute decision-making problem with the attribute value given in the form of hesitant fuzzy information has attracted the attention of scholars home and abroad. Xu [3–5] researched the distance measures, similarity degree, entropy and cross entropy in hesitant fuzzy sets, and apply them to multi-attribute decision-making. Farhadinia [6] discussed the relationship among hesitant fuzzy entropy, distance and similarity degree, and apply similarity degree into cluster analysis. In order to consider the decision maker's familiarity in professional field, Xia and Xu [7] proposed a credibility induced hesitant fuzzy aggregation operators, and applied it to the supplier selection.

Aiming at the fuzzy comprehensive evaluation problem of partial unknown attribute (index) weights information in alternatives, Xu [8] utilized IFHG operator to aggregate decision-making information and established the linear programming model, solved the corresponding index weight under scheme to maximize the score function, and established a mathematical programming model of the overall scheme of maximizing the score; Park [9] defined the correlation coefficient of interval-valued

J. Yang · C. Ruan (✉)
China School of Business Administration, South China University of Technology, Guangzhou 510641, China
e-mail: ruancyang@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_9

intuitionistic fuzzy number, and put forward a decision-making method based on interval-valued correlation coefficient decision-making method. Aiming at the fuzzy comprehensive evaluation problem of completely unknown attribute weights, Xu [10] defined the deviation degree between two intuitionistic fuzzy values, aggregated the individual intuitionistic fuzzy decision-making matrix into a comprehensive intuitionistic fuzzy decision-making matrix by using IFHA operators, and established the optimization model based on total deviation degree maximization to solve the optimal attribute weights; Liu [11] assign corresponding credibility to each evaluation values, gave several hesitant fuzzy information correlation formulas based on credibility, and established attribute weights model based on the correlation between evaluation matrix and preference of decision makers Ruan [12] proposes the weighted mutation rate correction incompletion G1 method which is used to measure the development balance degree and determines weights of evaluation indicators, which avoids the problem of integration of subjective and objective weights.

According to current literatures analysis, this paper attempts to study the following problems: how to take the credibility and the importance of hesitant fuzzy number weights into account. So we propose several correlation formulas considering the credibility, and construct several optimal attribute weights models based on the correlation between attribute evaluated values and the ideal values.

2 Basic Theory of Hesitant Fuzzy Sets

Zadeh [13] attempted to use membership function to express uncertain information firstly. However there exists lots of uncertainty in the actual decision-making problems, which leads to experts' hesitation in giving evaluation values. Hesitant fuzzy sets can solve this kind of problems effectively.

Definition 1 [1, 2] Let $X = \{x_1, x_2, \ldots, x_n\}$ be an non-empty set, then call

$$E = \{\langle x, h_E(x)\rangle | x \in X\} \tag{1}$$

the Hesitant Fuzzy Set. Among them $h_E(x)$ is a set of several possible membership values in [0, 1], which $x \in X$ means the membership grade of E to a set.

As to some practical problems such as the foundation and paper review, the familiarity with this field (credibility), usually need to be considered during the selection of evaluation experts, making the results more reasonable and credible. At present, when aggregating the hesitant fuzzy information, the decision-makers' familiarity with the professional field has been seldom considered, then Xia and Xu [7] put forward a hesitant fuzzy aggregation operator based on credibility.

Definition 2 [7] Let the hesitant fuzzy set be $h_1, h_2, \ldots, h_n, \forall \gamma_i \in h_i$, credibility $l_i \in [0, 1]$, and the weight vector of $h_i(1 = 1, 2, \ldots, n)$ is $w = (w_1, w_2, \ldots, w_n)^T$, $w_i \in [0, 1]$, $\sum_{i=1}^{n} w_i = 1$, then

$$CIHFWA(h_1, h_2, \ldots, h_n) = \bigoplus_{i=1}^{n} w_i(l_i h_i) = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2, \ldots, \gamma_n \in h_n}(1 - \prod_{i=1}^{n}(1 - l_i\gamma_i)^{w_i}) \tag{2}$$

is called the Credibility Induced Hesitant Fuzzy Weighted Averaging operator. When all credibility $l_i = 1$, the credibility induced hesitant fuzzy weighted operator degrades to the hesitant fuzzy weighted averaging operator.

$$HFWA(h_1, h_2, \ldots, h_n) = \bigoplus_{j=1}^{n} w_j h_j \tag{3}$$

Definition 3 [7] Let the hesitant fuzzy set be h_1, h_2, and its correlation degree satisfies the conditions as follows:

(1) $0 \le |C(h_1, h_2)| \le 1$
(2) $C(h_1, h_2) = 1$
(3) $C(h_1, h_2) = C(h_2, h_1)$

For that the elements number in may be different, in order to carry out the effective operation, elements should be added in the hesitant fuzzy set with less elements until the elements number in the set reaches $k = \max(k_1, k_2), k_1, k_2$ represent the elements number in the hesitant fuzzy set respectively. The adding principle can reflect the decision-makers' risk preference, and the decision-maker that prefers risk will add the element of the largest set value, while the decision-maker that hate risk will be on the contrary.

Based on the Definition 3, the literature [4] gave the following correlation degree formula about the hesitant fuzzy set:

$$C_1(h_1, h_2) = \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{\max\{\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2, \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2\}} \tag{4}$$

$$C_2(h_1, h_2) = \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{(\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 \cdot \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)^{1/2}} \tag{5}$$

Among which, $\gamma_1^{\sigma(j)}, \gamma_2^{\sigma(j)}$ denote the jth smallest element in the hesitant fuzzy set h_1, h_2 respectively, representing the elements number in hesitant fuzzy sets h_1, h_2.

Give another correlation degree formula below:

$$C_3(h_1, h_2) = \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{(\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 + \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)/2} \tag{6}$$

The above correlation degree formula (4)–(6) have the relationship as follows.

Property 1 *For any two hesitant fuzzy sets* h_1, h_2*, the hesitant fuzzy correlation degree formula* $C_i(h_1, h_2), i = 1, 2, 3$ *satisfies the following relationship:*

$$C_2(h_1, h_2) \geq C_3(h_1, h_2) \geq C_1(h_1, h_2) \tag{7}$$

Proof From the in equation $\sqrt{ab} \leq \frac{a+b}{2} \leq \max\{a, b\}, a, b \geq 0,$ we can get

$$(\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 \cdot \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)^{1/2} \leq (\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 + \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)/2 \leq \max\{\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2, \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2\}$$

then

$$\frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{(\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 \cdot \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)^{1/2}} \geq \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{(\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2 + \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2)/2} \geq \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{\max\{\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2, \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2\}}$$

Thus the conclusion is proved, and the proof is completed.

3 Determination Method of the Attribute Weight Based on Correlation Degree

In the multi-attribute comprehensive evaluation, sometimes the experts group need to give the weight of selected attributes. However, because of time pressure, the lack of knowledge and data, and the decision-makers' specialty limitation, the attribute weight information provided by the expert group tends to be incomplete. Therefore, how to carry out the schemes sorting and optimization according to the known evaluation knowledge is an interesting and important question. In this paper, we studied this kind of problem in the hesitant fuzzy environment, and put forward a kind of solution based on the correlation degree.

3.1 The Correlation Degree of the Hesitant Fuzzy Set and Its Properties Considering the Credibility

Considering the importance of the credibility, the following words will propose several hesitant fuzzy correlation formula based on the relative credibility and absolute credibility on the condition of Definition 3, and discuss their relationship.

Definition 4 [11] For any two hesitant fuzzy sets $h(x_1), h(x_2), k = \max(k_1, k_2)$, among which k_1, k_2 denotes the elements number in $h(x_1), h(x_2)$ respectively. $\forall \gamma_i \in h(x_i)$ there is credibility $l_i = [0, 1], i = 1, 2$, then we call

$$C_l^1 = \frac{\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})(l_2^{\tau(j)} \gamma_2^{\tau(j)})}{\max\{\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})^2, \sum_{j=1}^{k} (l_2^{\tau(j)} \gamma_2^{\tau(j)})^2\}} \tag{8}$$

the hesitant fuzzy correlation degree based on credibility. In which $\gamma_i^{\tau(j)} (j = 1, 2, \ldots, k)$ means the jth smallest element in $h(x_i)(i = 1, 2)$, and $l_i^{\tau(j)}$ means the corresponding absolute credibility to $\gamma_i^{\tau(j)}$.

Similarly, we can put forward the hesitant fuzzy correlation formula based on the absolute credibility as follows:

$$C_l^2 = \frac{\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})(l_2^{\tau(j)} \gamma_2^{\tau(j)})}{(\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})^2 \cdot \sum_{j=1}^{k} (l_2^{\tau(j)} \gamma_2^{\tau(j)})^2\})^{1/2}} \tag{9}$$

$$C_l^3 = \frac{\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})(l_2^{\tau(j)} \gamma_2^{\tau(j)})}{(\sum_{j=1}^{k} (l_1^{\tau(j)} \gamma_1^{\tau(j)})^2 + \sum_{j=1}^{k} (l_2^{\tau(j)} \gamma_2^{\tau(j)})^2)/2} \tag{10}$$

Property 2 *For any two hesitant fuzzy set, the hesitant fuzzy correlation degree satisfies the following relationship:*

$$C_l^2(h_1, h_2) \geq C_l^3(h_1, h_2) \geq C_l^1(h_1, h_2) \tag{11}$$

Definition 5 For any two hesitant fuzzy sets $h(x_1), h(x_2), k = \max(k_1, k_2)$, among which k_1, k_2 denotes the elements number in $h(x_1), h(x_2)$ respectively. $\forall \gamma_i \in h(x_i)$ there is credibility $l_i = [0, 1], i = 1, 2$, then we call

$$C_r^1 = \frac{\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})(l_2^{(j)} \gamma_2^{\tau(j)})}{\max\{\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})^2, \sum_{j=1}^{k} (l_2^{(j)} \gamma_2^{\tau(j)})^2\}} \tag{12}$$

the hesitant fuzzy correlation degree based on credibility. In which $\gamma_i^{\tau(j)} (j = 1, 2, \ldots, k)$ means the jth smallest element in $h(x_i)(i = 1, 2)$, $l_i^{\tau(j)}$ means the corresponding absolute credibility to $\gamma_i^{\tau(j)}$, and $l_i^{(j)} = n \times \frac{l_i^{\tau(j)}}{\sum_{i=1}^{k} l_i^{\tau(j)}}$ means the corresponding relative credibility to $\gamma_i^{\tau(j)}$. Similarly, we can put forward the hesitant fuzzy correlation formula based on the absolute credibility as follows:

$$C_r^2 = \frac{\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})(l_2^{(j)} \gamma_2^{\tau(j)})}{(\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})^2 \cdot \sum_{j=1}^{k} (l_2^{(j)} \gamma_2^{\tau(j)})^2\})^{1/2}} \tag{13}$$

$$C_r^3 = \frac{\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})(l_2^{(j)} \gamma_2^{\tau(j)})}{(\sum_{j=1}^{k} (l_1^{(j)} \gamma_1^{\tau(j)})^2 + \sum_{j=1}^{k} (l_2^{(j)} \gamma_2^{\tau(j)})^2)/2} \tag{14}$$

Property 3 *For any two hesitant fuzzy set, the hesitant fuzzy correlation degree satisfies the following relationship:*

$$C_r^2(h_1, h_2) \geq C_r^3(h_1, h_2) \geq C_r^1(h_1, h_2) \tag{15}$$

The selection of large, geometric mean and arithmetic mean in the correlation formula reflect the decision-maker's risk attitude, which is risk preference, disgust and neutral. The decision-makers can choose different correlation formula based on actual needs.

3.2 Determination of the Attribute Weight

Consider the multi-index comprehensive evaluation problem in the hesitant fuzzy environment. Let $A = \{A_1, A_2, \ldots, A_m\}$ be the scheme set, $G = \{G_1, G_2, \ldots, G_n\}$ be the index set. $a_{ij}(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ denotes the aggregation value of the jth attribute in the ith scheme, $w = (w_1, w_2, \ldots, w_n)$ is the weight vector of the attribute, $w_j \in [0, 1]$, $\sum_{j=1}^{n} w_j = 1$, $w^* = (w_1^*, w_2^*, \ldots, w_n^*)$ is the optimal attribute weight vector, $w_j^* \in [0, 1]$, $\sum_{j=1}^{n} w_j^* = 1$. H denotes the partial weight information proposed by the expert group.

Method 1: If the decision-making group has no preference for the alternatives, the scheme preference is not considered.

Suppose $C_1(A_i, A^*)(i = 1, 2, \ldots, m)$ is the correlation between the scheme $A_i(i = 1, 2, \ldots, m)$ and the ideal value A^*. The definition of is in 2.1, and $A^* = (a_1^*, a_2^*, \ldots, a_n^*)$ is the set of the maximum attribute value in each scheme aggregating from CIHFHA. In general, the actual evaluation value of the scheme $A_i = (i = 1, 2, \ldots, m)$ has a certain gap with the ideal value A^*, so the larger the value of $C_1(A_i, A^*)$, the more close A_i is to A^*, which means the better the scheme A_i is. Therefore, the reasonable weight should make $C_1(A_i, A^*)(i = 1, 2, \ldots, m)$ as large as possible. Based on this, we can construct the optimization model to solve the weight.

$$\textbf{Model 1} \quad \max\left(\sum_{j=1}^{n} w_j C_1(a_{1j}, a_j^*), \sum_{j=1}^{n} w_j C_1(a_{2j}, a_j^*), \ldots, \sum_{j=1}^{n} w_j C_1(a_{nj}, a_j^*)\right) \quad (16)$$

Each scheme is coequal, so the scheme preference does not exist. Therefore, we can aggregate the above optimization model to the single objective optimization model according to the weight [14]:

$$\textbf{Model 2} \quad \begin{array}{c} \max \sum_{i=1}^{m} \sum_{j=1}^{n} w_j C_1(a_{ij}, a_j^*) \\ s.t. w \in H \end{array} \quad (17)$$

Model 2 can be easily solved by the single shape method used in the linear programming model, and we can get the optimal solution $w^* = (w_1^*, w_2^*, \ldots, w_n^*)$ as the index weight vector.

If the weight vector is unknown, we can utilize the following model to obtain the weight vector:

Model 3
$$\max \sum_{i=1}^{m} \sum_{j=1}^{n} w_j C_1(a_{ij}, a_j^*) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_j \frac{\sum_{j=1}^{l} \gamma_1^{\sigma(j)} \cdot \gamma_2^{\sigma(j)}}{\max\{\sum_{j=1}^{l} (\gamma_1^{\sigma(j)})^2, \sum_{j=1}^{l} (\gamma_2^{\sigma(j)})^2\}}$$
$$\text{s.t.} \sum_{j=1}^{n} w_j^2 = 1, w_j \geq 0, j = 1, 2, \ldots, n \tag{18}$$

Without loss of generality, the correlation degree formula $C_1(a_{ij}, a_j^*)$ in model 3 is calculated according to formula (4). To solve the equation, we construct the Lagrange Function:

$$L(w, \lambda) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_j C_1(a_{ij}, a_j^*) + \frac{\lambda}{2}(\sum_{j=1}^{n} w_j^2 - 1) \tag{19}$$

Calculating the partial derivative, and let them be 0. Then we have

$$\begin{cases} \frac{\delta L(w_j, \lambda)}{\delta w_j} = \sum_{i=1}^{m} C_1(a_{ij}, a_j^*) + \lambda w_j = 0 \\ \frac{\delta L(w_j, \lambda)}{\delta \lambda} = \sum_{j=1}^{n} w_j^2 - 1 = 0 \end{cases} \tag{20}$$

Solve the equations, and we have

$$w_j = \frac{\sum_{i=1}^{m} C_1(a_{ij}, a_j^*)}{\sqrt{\sum_{j=1}^{n} (\sum_{i=1}^{m} C_1(a_{ij}, a_j^*))^2}} \tag{21}$$

Processing w_j to unitization and the optimal attribute weight is

$$w_j^* = \frac{\sum_{i=1}^{m} C_1(a_{ij}, a_j^*)}{\sqrt{\sum_{j=1}^{n} (\sum_{i=1}^{m} C_1(a_{ij}, a_j^*))}} \tag{22}$$

Substitute the correlation formula (4) to the above formula, and we have

$$w_j^* = \frac{\sum_{i=1}^{m} \frac{\sum_{j=1}^{l} a_{ij} \cdot a_j^*}{\max\{\sum_{j=1}^{l} (a_{ij})^2, \sum_{j=1}^{l} (a_j^*)^2\}}}{\sum_{j=1}^{n} (\sum_{i=1}^{m} \frac{\sum_{j=1}^{l} a_{ij} \cdot a_j^*}{\max\{\sum_{j=1}^{l} (a_{ij})^2, \sum_{j=1}^{l} (a_j^*)^2\}})}, j = 1, 2, \ldots, n \tag{23}$$

In particular, when I = 1, and the ideal values are the maximum value of each attribute comprehensive values. Then

$$w_j^* = \frac{\sum_{i=1}^{m} \frac{a_{ij}}{a_j^*}}{\sum_{j=1}^{n} (\sum_{i=1}^{m} \frac{a_{ij}}{a_j^*})}, j = 1, 2, \ldots, n \tag{24}$$

w_j^* is the optimal attribute weight vector solved by optimization model constructed according to the correlation degree.

Method 2: Determination of the Attribute Weight Considering the Scheme Preference

For the multi-attribute decision-making problems with credibility and the attribute value from the hesitant fuzzy set, we propose the hesitant fuzzy information attribute weight determination model with scheme preference, considering the decision-maker's subjective preference and objective preference. The decision-making group gives the preference value $\overset{\Delta}{h}_i (i = 1, 2, \ldots, m)$ for the scheme, and measure the scheme according to the attributes, and the hesitant fuzzy decision-making matrix $D_l = (h_{ijk})_{m \times n \times p}$ with credibility is worked out, h_{ijk} denoting the kth hesitant fuzzy number of the jth attribute in the ith scheme. The attribute value h_{ij} in the scheme $D_l = (h_{ij})_{m \times n}$ can be seen as objective preference, and the preference value $\overset{\Delta}{h}_i (i = 1, 2, \ldots, m)$ can be seen as the subjective preference. Because of the influence of the decision-makers and decision-making environment, there exists certain correlation between the subjective and objective preferences. In order to make the decision-making more reasonable, the selection of the attribute weight vector $w = (w_1, w_2, \ldots, w_n)$ should make the overall correlation degree between the subjective and objective preferences the biggest. Therefore, we propose the hesitant fuzzy attribute weight determination model considering the credibility [11, 13]:

Model 4

$$\max \sum_{j=1}^{n} C_r^1(h_{ij}, \overset{\Delta}{h}_i) w_t = \sum_{j=1}^{n} \frac{\sum_{k=1}^{p} (l_{ij}^{(k)} \gamma_{ij}^{\tau(k)})(l_{ij}^{\overset{\Delta}{(k)}} \gamma_{ij}^{\overset{\Delta}{\tau}(k)})}{\max\{\sum_{j=1}^{k} (l_{ij}^{(k)} \gamma_{ij}^{\tau(k)})^2, \sum_{j=1}^{k} (l_{ij}^{\overset{\Delta}{(k)}} \gamma_{ij}^{\overset{\Delta}{\tau}(k)})^2\}} w_j \tag{25}$$

$$\text{S.t.} \sum_{j=1}^{n} w_j^2 = 1, w_j \geq 0, j = 1, 2, \ldots, n$$

Without loss of generality, the correlation degree formula $C_r^1(h_{ij}, \overset{\Delta}{h}_i)$ in model 4 is calculated according to formula (12), reflecting the correlation between the subjective and objective preference of decision-makers. The schemes are competing fairly, and no preference relationship, so model 4 can turn to the single objective planning model:

Model 5

$$\max \sum_{i=1}^{m} \sum_{j=1}^{n} C_r^1(h_{ij}, \overset{\Delta}{h}_i) w_t = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\sum_{k=1}^{p} (l_{ij}^{(k)} \gamma_{ij}^{\tau(k)})(l_{ij}^{\overset{\Delta}{(k)}} \gamma_{ij}^{\overset{\Delta}{\tau}(k)})}{\max\{\sum_{j=1}^{k} (l_{ij}^{(k)} \gamma_{ij}^{\tau(k)})^2, \sum_{j=1}^{k} (l_{ij}^{\overset{\Delta}{(k)}} \gamma_{ij}^{\overset{\Delta}{\tau}(k)})^2\}} w_j$$

$$\text{s.t.} \sum_{j=1}^{n} w_j^2 = 1, w_j \geq 0, j = 1, 2, \ldots, n \tag{26}$$

Then we have

$$w_j^* = \frac{\sum_{i=1}^{m} C_r^1(a_{ij}, a_j^*)}{\sum_{j=1}^{n} (\sum_{i=1}^{m} C_r^1(a_{ij}, a_j^*))} \tag{27}$$

4 Concluding Remarks

In order to get more reasonable result, we give the corresponding credibility to represent the decision-maker's familiarity with the field. This paper proposed several kinds of optimization model to solve the attribute weight based on the correlation degree considering the credibility, which could utilize the known information effectively, and the calculated attribute weight would be more subjective and reasonable. So the method has high practical value, and can be applied to personnel management, supplier selection, evaluation of economic efficiency etc.

References

1. Torra, V.: Hesitant fuzzy sets [J]. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
2. Torra, V., Narukawa, Y.: On hesitant fuzzy sets and decision. In: Proceedings of the 18th IEEE International Conference on Fuzzy systems, pp. 1378–1382 (2009)
3. Xu, Z.S., Xia, M.M.: Distance and similarity measures for hesitant fuzzy sets. Inf. Sci. **181**(11), 2128–2138 (2011)
4. Xu, Z.S., Xia, M.M.: On distance and correlation measures of hesitant fuzzy information. Int. J. Intell. Syst. **26**(5), 410–425 (2011)
5. Xu, Z.S., Xia, M.M.: Hesitant fuzzy entropy and cross-entropy and their use in multiattribute decision-making. Int. J. Intell. Syst. **27**(9), 799–822 (2012)
6. Farhadinia, B.: Information measures for hesitant fuzzy sets and interval-valued hesitant fuzzy sets. Inf. Sci. **240**(10), 129–144 (2013)
7. Xia, M.M., Xu, Z.S., Chen, N.: Induced aggregation under confidence levels. Int. J. Uncertainty, Fuzziness Knowl-Based Syst. **19**(2), 201–227 (2011)
8. Xu, Z.S.: Multi-person multi-attribute decision making models under intuitionistic fuzzy environment. Fuzzy Optim. Decis. Making **6**(3), 221–236 (2007)
9. Park, D.G., Kwun, Y.C., Park, J.H.: Correlation coefficient of interval-valued intuitionistic fuzzy sets and its application to multiple attribute group decision making problems. Math. Comput. Model. **50**(9–10), 1279–1293 (2009)
10. Xu, Z.S.: A deviation-based approach to intuitionistic fuzzy multiple attribute group decision making. Group Decis. Negot. **19**(1), 57–76 (2010)
11. Liu, X.D., Zu, J.J, Zhang, S.T.: Hesitant fuzzy decision making method with confidence levels and preference information on alternatives. Syst. Eng. Electron. 1–7 (2014)
12. Ruan, C.Y., Yang, J.H.: Software quality evaluation model based on weighted mutation rate correction incompletion G1 combination weights. Math. Probl. Eng. 2014. (2014)
13. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
14. Zeng, S.Z.: Study on the comprehensiv evaluation technology based on intuitionistic fuzzy information. Hangzhou: Doctoral Dissertation of Zhejiang Gongshang University (2013)

Pan-uncertain Measure

Chao Wang, Minghu Ha and Xiaowei Liu

Abstract Probability measure, possibility measure and uncertain measure can effectively deal with random phenomenons, fuzzy phenomenons, uncertainty phenomenons in real-world correspondingly. In order to establish a unified tool to handle random phenomenons, fuzzy phenomenons and uncertainty phenomenons, pan-uncertain measure is defined by normality, pan duality, pan subadditivity. Then, corresponding pan-uncertain product measure and pan-uncertainty space are given. In the end, pan-uncertain variable and pan-uncertain distribution are defined and discussed on pan-uncertain space.

Keywords Pan-uncertain measure · Pan-uncertain variable · Pan-uncertain distribution · Uncertain measure

1 Introduction

The probability measure formulated axiomatically by Kolmogorov [1] in 1933 is a non-negative, norm, countable additive set function, which is an effective tool to deal with the random phenomenon in the real-world. While the additivity property is too restrictive in some application contexts, that there exist more non-additive set functions [2–9]. For example, possibility measure as an important non-additive measure, emerged from the concept of fuzzy set [10], which was proposed by Zadeh [2, 3] via membership function. It can measure the fuzzy events and process the fuzzy phenomenon effectively.

C. Wang (✉) · M. Ha
School of Economics and Management, Hebei University of Engineering, Handan 056038, People's Republic of China
e-mail: wang222chao@163.com

X. Liu
School of Science, Hebei University of Engineering, Handan 056038, People's Republic of China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_10

In addition to the random and fuzzy phenomenon, there also exist uncertainty phenomenons in the real-world. Liu [9] wrote that “when the sample size is too small (even no-sample) to estimate a probability distribution, we have to invite some domain experts to evaluate their belief degree that each event will occur. Since human beings usually overweight unlikely events, the belief degree may have much larger variance than the real frequency, this phenomenon was named uncertainty.” To deal with the above uncertainty phenomenon, uncertain measure and corresponding uncertainty theory were founded by Liu [8] in 2007. Subsequently, uncertainty theory has been applied into many aspects [11, 12].

Uncertain measure is a non-additive set function with normality, self-duality, countable subadditivity, and product axioms. Both possibility measure with normality, fuzzy additivity and probability measure with normality, monotonicity, countable additivity satisfy the first three axioms of uncertain measure, but probability measure cannot be regarded as a special case of uncertain measure because it does not satisfy the product axiom. The operator of probability product measure is multiplication ‘$\cdot$’, while the operator of uncertain product measure is minimization ‘$\wedge$’. On the other hand, though random, fuzzy and uncertainty are three different phenomenons, there are also some relations. When the sample size is too small to estimate a probability distribution, the uncertainty appears, we should deal with uncertainty theory [9]. When the sample size becomes large, the uncertainty disappears. Then, the problem becomes probabilistic (random). Besides, fuzzy and random can make up fuzzy random phenomenons, and uncertainty and random can also make up uncertain random phenomenons. Therefore, it is important to establish a unified tool to process the above three phenomenons. Then, pan-uncertain measure is proposed with normality, pan duality, pan subadditivity, and corresponding pan-uncertain theory is initially founded, which can deal with random, fuzzy, and uncertainty phenomenons.

2 Preliminary

Let be Ω a nonempty set, and be $\mathcal{A}$ a σ-algebra over Ω. Each element $A \in \mathcal{A}$ in called an event, and $P(A)$ indicates the probability that A may occur. Probability measure [1] P is a set function satisfying the following three axioms:
Axiom 1. $P(\Omega) = 1$.
Axiom 2. $P(A) \geq 0, \forall A \in \mathcal{A}$.
Axiom 3. For every countable sequence of mutually disjoint events $\{A_i\}, A_i \in \mathcal{A}$, we have

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right).$$

The triples $(\Omega, \mathcal{A}, P)$ is called a probability measure space. Let $(\Omega_k, \mathcal{A}_k, P_k), k = 1, 2, \ldots$ be probability measure spaces. The probability product measure is defined by

$$P\left(\prod_{k=1}^{\infty} A_k\right) = \prod_{k=1}^{\infty} P_k\left(A_k\right)$$

where A_k are chosen from $\mathcal{A}_k, k = 1, 2, \ldots$

Zadeh [2, 3] introduce possibility measure to deal with fuzzy problems. It satisfies the following three axioms:
Axiom 1. $Pos(\Omega) = 1$ for the universal set Ω.
Axiom 2. $Pos(\phi) = 0$ for empty set ϕ.
Axiom 3. For every sequence events $\{A_i\}$, we have

$$Pos\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigvee_{\mathrm{i}=1}^{\infty} Pos\left(A_i\right).$$

In order to deal with human uncertainty, Liu [8] proposed uncertain measure in 2007. Let Γ be a nonempty set, and $\mathcal{L}$ be a σ-algebra over Γ. Each element $\Lambda \in \Omega$ called an uncertain event, uncertain measure is interpreted as the personal belief degree of the uncertain event Λ. Then, the uncertain measure is defined by
Axiom 1. $\mathcal{M}(\Gamma) = 1$ for the universal set Γ.
Axiom 2. $\mathcal{M}(\Lambda) + \mathcal{M}(\Lambda^c)$ for any even $\Lambda \in \mathcal{L}$.
Axiom 3. For every countable sequence events $\{\Lambda_i\}$, we have

$$\mathcal{M}\left(\bigcup_{i=1}^{\infty} \Lambda_i\right) \leq \sum_{i=1}^{\infty} \mathcal{M}\left(\Lambda_i\right).$$

Let $\left(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k\right), k = 1, 2, \ldots$ be uncertain spaces. Then, the uncertain product measure $\mathcal{M}$ is given by

$$\mathcal{M}\left(\prod_{k=1}^{\infty} \Lambda_k\right) = \bigwedge_{k=1}^{\infty} \mathcal{M}_k\left(\Lambda_k\right)$$

for each element $\Lambda_k \in \Gamma_k$.

Pan-uncertain Measure

In order to establish a unified tool to process random phenomenons, fuzzy phenomenons and uncertainty phenomenons in the real-world, we first proposed pan-uncertain measure based on two binary pan operators [13].

Definition 1 Let $\oplus$ and $\odot$ be two binary pan operators on $[0, \infty]$. If for any $a, b, c, a_i, b_i \in [0, \infty](i = 1, 2)$, we have

(1) $a \oplus b = b \oplus a, a \odot b = b \odot a$.
(2) $(a \oplus b) \oplus c = a \oplus (b \oplus c), (a \odot b) \odot c = a \odot (b \odot c)$.
(3) $a_1 \leq b_1, a_2 \leq b_2 \Rightarrow a_1 \oplus a_2 \leq b_1 \oplus b_2, a_1 \odot a_2 \leq b_1 \odot b_2$.

(4) There exists a zero element $\mathbf{0}$, and $a \oplus \mathbf{0} = a; a \odot \mathbf{0} = \mathbf{0}$.
(5) There exists an unit element $\mathbf{I}$, and $a \odot \mathbf{I} = \mathbf{I} \odot a = a$.
(6) If $\{a_n\} \subset [0, \infty], \{b_n\} \subset [0, \infty]$ and $\lim_{n\to\infty} a_n < \infty$, $\lim_{n\to\infty} b_n < \infty$. Then

$$\lim_{n\to\infty} \left(a_n \oplus b_n\right) = \lim_{n\to\infty} a_n \oplus \lim_{n\to\infty} b_n, \lim_{n\to\infty} \left(a_n \odot b_n\right) = \lim_{n\to\infty} a_n \odot \lim_{n\to\infty} b_n.$$

It is easy to find that '$\cdot$' and '$\wedge$' satisfy the pan operator '$\odot$', '$+$' and '$\vee$' satisfy the pan operator '$\oplus$'.

In the paper, let X be a nonempty set, and $\mathcal{F}$ be a σ-algebra over X. Each element $A \in \mathcal{F}$ called a pan-uncertain event.

Definition 2 A pan-uncertain measure μ is a mapping from $\mathcal{F}$ to $[0, \infty]$ satisfying
Axiom 1. $\mu(\phi) = \mathbf{0}$ for the empty set ϕ, where $\mathbf{0}$ is a zero element.
Axiom 2. $\mu(A) \oplus \mu(A^c) = \mathbf{I}$ for any pan-uncertain event $A \in \mathcal{F}$, where $\mathbf{I}$ is an unit element.
Axiom 3. For every countable sequence pan-uncertain events A_i, we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \bigoplus_{i=1}^{\infty} \mu\left(A_i\right).$$

In this paper, $(X, \mathcal{F}, \mu)$ is called a pan-uncertain space.

Let $(X_k, \mathcal{F}_k, \mu_k), k = 1, 2, \ldots$ be pan-uncertain spaces. Then, the pan-uncertain product measure is defined by

$$\mu\left(\prod_{k=1}^{\infty} A_k\right) = \bigodot_{k=1}^{\infty} \mu_k\left(A_k\right) \tag{1}$$

Let $A, B \in \mathcal{F}$ and $A \subset B$. Then, the pan-uncertain measure is monotonic if

$$\mu(A) \leq \mu(B).$$

Property 1 $\mu(X) = \boldsymbol{I}$ *for the universal set* X.

Proof $\mathbf{I} = \mu(X) \oplus \mu(\phi) = \mu(X) \oplus \mathbf{0} = \mu(X)$

Property 2 $\mu(A) \leq \boldsymbol{I}$ *for any pan-uncertain event* A

Proof As $\mu(A^c) \geq 0$, we have $\mu(A) = \mu(A) \oplus 0 \leq \mu(A) + \mu(A^c) = \mathbf{I}$

Property 3 $\mu(A_1 \cup A_2) \leq \mu(A_1) \oplus \mu(A_2)$ *for any Pan-uncertain events* A_1, A_2.

Proof Denote $A_3 = A_4 = \cdots = \phi$. As $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \bigoplus_{i=1}^{\infty} \mu\left(A_i\right)$, we have

$$\mu\left(A_1 \cup A_2\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \le \bigoplus_{i=1}^{\infty} \mu\left(A_i\right) = \mu\left(A_1\right) \oplus \mu\left(A_2\right) \oplus \left(\bigoplus_{i=3}^{\infty} \mu\left(A_i\right)\right)$$
$$= \mu\left(A_1\right) \oplus \mu\left(A_2\right).$$

Theorem 1 *Let the operator '⊕' be '+'. Then, the pan-uncertain measure μ satisfies monotonicity.*

Proof If $A \subset B$, we have $A^c \cup B = X$. Then,

$$\mathbf{I} = \mu(X) = \mu(A^c \cup B) \le \mu(A^c) + \mu(B) = I - \mu(A) + \mu(B) \Rightarrow \mu(A) \le \mu(B)$$

Remark 1 When the operator '⊕' is '∨', the pan-uncertain measure μ does not satisfy monotonicity sometimes. For example, let $X = \left\{x_1, x_2, x_3\right\}$. A pan-uncertain measure μ is given by

$$\mu\left\{x_1\right\} = 0.4, \mu\left\{x_2\right\} = 0.8, \mu\left\{x_3\right\} = 1$$
$$\mu\left\{x_1, x_2\right\} = 0.7, \mu\left\{x_1, x_3\right\} = 1, \mu\left\{x_2, x_3\right\} = 1$$
$$\mu\left\{x_1, x_2, x_3\right\} = 1, \mu\left\{\phi\right\} = 0$$

It is obvious that

$$\{x_2\} \subset \{x_1, x_2\}, \textit{while } \mu\{x_2\} > \mu\{x_1, x_2\}$$

Theorem 2 *Let the operator '⊕' be '∨' and **I** = 1. Then, the monotonic pan-uncertain measure μ is also a possibility measure.*

Proof It is obvious that the zero element $\mathbf{0} = 0$. Then we have

$$\mu(\phi) = 0, \mu(X) = 1.$$

Since $A_i \subset \bigcup_{i=1}^{\infty} A_i, i = 1, 2, \ldots,$ we have

$$\mu\left(A_i\right) \le \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \Rightarrow \bigvee_{i=1}^{\infty} \mu\left(A_i\right) \le \mu\left(\bigcup_{i=1}^{\infty} A_i\right), i = 1, 2, \ldots.$$

For every countable sequence pan-uncertain events $\{A_i\}$, we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \le \bigvee_{i=1}^{\infty} \mu\left(A_i\right),$$

by the use of the Axiom 3 of pan-uncertain measure. Then,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigvee_{\mathrm{i}=1}^{\infty} \mu\left(A_i\right).$$

i.e. the pan-uncertain measure satisfies the three axioms of possibility measure.

Remark 2 It is obvious that the pan-uncertain measure is also a uncertain measure while the operator '$\oplus$' is '+', the operator '$\odot$' is '$\wedge$' and $\mathbf{I} = 1$

For the operator '$\odot$', both probability product measure and uncertainty product measure satisfy the Eq. (1). It is that probability measure, possibility measure, and uncertain measure are all the special case of pan-uncertain measure. Therefore, the pan-uncertain measure indicates the pan level (probability, possibility, and uncertainty) that pan-uncertain even will occur.

However, some pan-uncertain measure does not satisfy monotonicity such as the example in Remark 1, and it is not probability measure, possibility measure or uncertain measure. Since probability measure, possibility measure, and uncertain measure are all monotonic, we mainly study the monotonic pan-uncertain measure in the rest of the paper.

Theorem 3 *Let μ be a monotonic pan-uncertain measure. Then*

(1) $\mu(A_1) \oplus \mu(A_2) \le \boldsymbol{I}$ *for any disjoint events A_1 and A_2;*
(2) $\lim_{i\to\infty} \mu\left(A \cup A_i\right) = \lim_{i\to\infty} \mu\left(A \backslash A_i\right) = \mu(A)$ *for every countable sequence pan-uncertain events A_i with* $\lim_{i\to\infty} \mu\left(A_i\right) = 0$.

Proof (1) Since $A_1 \cap A_2 = \phi \Rightarrow A_1 \subset A_2^c$, we have

$$\mu\left(A_1\right) \oplus \mu\left(A_2\right) \le \mu\left(A_2^c\right) \oplus \mu\left(A_2\right) = \mathbf{I}$$

(2) As $\mu(A) \le \mu\left(A \cup A_i\right) \le \mu(A) \oplus \mu\left(A_i\right)$. Then

$$\mu(A) \le \lim_{i\to\infty} \mu\left(A \cup A_i\right) \le \mu(A) \oplus \lim_{i\to\infty} \mu\left(A_i\right) = \mu(A).$$

Since $A \backslash A_i \subset A = \left(A \backslash A_i\right) \cup A_i$, so

$$\lim_{i\to\infty} \mu\left(A \backslash A_i\right) \le \mu(A) \le \lim_{i\to\infty} \mu\left(A \backslash A_i\right) \oplus \lim_{i\to\infty} \mu\left(A_i\right) = \lim_{i\to\infty} \mu\left(A \backslash A_i\right).$$

3 Pan-uncertain Variables

Definition 3 Let μ be a monotonic pan-uncertain measure. Pan-uncertain variable is a measurable mapping from pan-uncertain space $(X, \mathcal{F}, \mu)$ to real number set R, i.e. for any Borel set B, the set $\{\omega \in X \mid \xi(\omega) \in B\}$ is a pan-uncertain event.

Definition 4 Let ξ, η be pan-uncertain variables on pan-uncertain space. Then

(1) If $\mu\{\xi < 0\} = 0$, the pan-uncertain variable ξ is nonnegative.
(2) If $\mu\{\xi \le 0\} = 0$, the pan-uncertain variable ξ is positive.
(3) If $\mu\{\xi \ne \eta\} = 0$, the pan-uncertain variables ξ, η are almost equal.

Definition 5 The pan-uncertain distribution of pan-uncertain variable ξ is defined by

$$\varphi(x) = \mu\{\xi \le x\}, x \in R$$

Property 4 *The pan-uncertain distribution of pan-uncertain variable ξ satisfies*

(1) $0 \le \varphi(x) \le \boldsymbol{I}$.
(2) $\varphi(x) \le \varphi(y), \forall x \le y \in R$.
(3) $\varphi(x) = 0 \Rightarrow \mu\{\xi > x\} = \boldsymbol{I}$*; While* $\varphi(x) = \boldsymbol{I}$, *we cannot get* $\mu\{\xi > x\} = 0$.

Proof It is obvious by the use of the properties of monotonic pan-uncertain measure.

Remark 3 If the pan-uncertain measure does not satisfy monotonicity, the corresponding distribution may not be monotonic (see Example 1). So, the pan-uncertain variable is defined based on monotonicity pan-uncertain measure.

Example 1 Let the binary operator '$\oplus$' be '$\vee$', $\mathbf{I} = 1$, and $X = \{x_1, x_2, x_3\}$. A pan-uncertain measure is given by

$$\mu\{x_1\} = 0.8, \mu\{x_2\} = 0.4, \mu\{x_3\} = 1,$$
$$\mu\{x_1, x_2\} = 0.7, \mu\{x_1, x_3\} = 1, \mu\{x_2, x_3\} = 1,$$
$$\mu\{x_1, x_2, x_3\} = 1, \mu\{\phi\} = 0.$$

Then the variable

$$\xi(\omega) = \begin{cases} 1, \ \omega = x_1 \\ 2, \ \omega = x_2 \\ 3, \ \omega = x_3 \end{cases}$$

has a non-monotonicity distribution as follows

$$\varphi(x) = \begin{cases} 0, & x < 1 \\ 0.8, & 1 \le x < 2 \\ 0.4, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}.$$

Definition 6 Let ξ, η be pan-uncertain variables on pan-uncertain space. Then

(1) the pan-uncertain variables ξ, η are identical distributed if

$$\mu\{\xi \le x\} = \mu\{\eta \le x\}, \forall x \in R.$$

(2) the pan-uncertain variables ξ, η are said to be independent if

$$\mu\{\xi \le x, \eta \le x\} = \mu\{\xi \le x\} \odot \mu\{\eta \le x\}, \forall x \in R.$$

Definition 7 Let $\xi_1, \xi_2, \ldots, \xi_n$ be pan-uncertain variables on pan-uncertain space. Then $(\xi_1, \xi_2, \ldots, \xi_n)$ is called an n-dimension uncertain variable.

Definition 8 Let $(\xi_1, \xi_2, \ldots, \xi_n)$ be an n-dimension uncertain variables. Then the joint uncertain distribution is defined by

$$\varphi\left(x_1, x_2, \ldots, x_n\right) = \mu\left(\bigcap_{i=1}^{n}\left\{\xi_i \le x_i\right\}\right), x_i \in R$$

Definition 9 The pan-uncertain variables $\xi_1, \xi_2, \ldots, \xi_n$ are said to be independent iff

$$\mu\left(\bigcap_{i=1}^{n}\left\{\xi_i \in B_i\right\}\right) = \bigodot_{i=1}^{n} \mu\left\{\xi_i \in B_i\right\}$$

for any Borel set B_i.

Theorem 4 *If $\xi_1, \xi_2, \ldots, \xi_n$ are independent pan-uncertain variables with pan-uncertain distributions $\varphi_1, \varphi_2, \ldots, \varphi_n$, respectively. Then, we have*

$$\varphi\left(x_1, x_2, \ldots, x_n\right) = \bigodot_{i=1}^{n} \varphi\left(x_i\right), x_i \in R$$

Proof It is obvious by the use of Definitions 8 and 9.

4 Conclusion

In this paper, pan-uncertain measure, pan-uncertain space, pan-uncertain variable, and pan-uncertain distribution are defined, and the properties of pan-uncertain measure and pan-uncertain variable are discussed. Then, the pan-uncertain theory is initially established, which covers the probability theory, possibility theory and uncertainty theory, and it can deal with random, fuzzy, and uncertainty phenomenons in the real-world.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No.61073121), the Natural Science Foundation of Hebei Province of China (No.F2015402033).

References

1. Kolmogorov, A.N.: Grundbegriffe der Wahrscheinlichkeitsrechnung. Julius Springer, Berlin (1933)
2. Zadeh, L.A.: Probability measures of fuzzy events. J. Math. Anal. Appl. **23**, 421–27 (1968)
3. Zadeh, L.A: Fuzzy set as a basis for a theory of possibility. Fuzzy Sets Syst. **1**, 3-28 (1978)
4. Wang, Z.Y., Klir, G.J.: Fuzzy Measure Theory. Plenum Press, New York (1992)
5. Pap, E.: Null-Additive Set Functions. Kluwer, Boston (1995)
6. Wang, Z.Y., Klir, G.J.: Generalized Measure Theory. Springer, New York (2008)
7. Ha, M.H., Yang, L.Z., Wu, C.X.: Introduction of Generalized Fuzzy Set-valued Measure. Science Press, Beijing (2009)
8. Liu, B.D.: Uncertainty Theory, 2nd edn. Springer, Berlin (2007)
9. Liu, B.D.: Uncertainty Theory (4th edn). UTLAB (2012)
10. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
11. Liu, B.D.: Theory and Practice of Uncertain Programming, 2nd edn. Springer, Berlin (2009)
12. Liu, B.D.: Uncertain risk analysis and uncertain reliability analysis. J. Uncertain Syst. **4**, 163–170 (2010)
13. Yang, Q.J.: Pan integral on fuzzy measure space. J. Hebei Univ. (Nature Science Edition) **2**, 102–106 (1984)

Fuzzy Risk Analysis Method Based on Trapezoidal Intuitionistic Fuzzy Numbers

Xiaojie Zhou, Shuai Wang and Cheng Zhang

Abstract In this paper, a new fuzzy risk analysis method is proposed based on trapezoidal intuitionistic fuzzy numbers. A new weighted similarity measure between trapezoidal intuitionistic fuzzy numbers is presented based on centers-of-gravity of membership and nonmembership functions. In a fuzzy risk analysis problem, the linguistic values and probability of failure are represented as trapezoidal intuitionistic fuzzy numbers. The proposed similarity is applied to translate the risk of failure into proper linguistic value. It is showed that the proposed model is an effective method for fuzzy risk analysis and the weight in similarity measure improves the flexibility of a system.

Keywords Fuzzy risk analysis · Trapezoidal intuitionistic fuzzy number · Similarity

1 Introduction

Risk analysis plays an important role in many decision problems, and many researches are focused on fuzzy risk analysis in recent years [1–4]. A natural way is to represent the probability of failure as a generalized fuzzy numbers and measure the risk by similarity. Chen and Chen proposed a simple center of gravity method to calculate the center-of-gravity points of generalized fuzzy numbers, and then the degree of similarity between two generalized fuzzy numbers was measured based on the center-of-gravity points [1]. By combining the concepts of geometric distance, the perimeter and the height of generalized fuzzy numbers, Wei and Chen presented a new similarity measure between generalized fuzzy numbers and used the new similarity to fuzzy risk analysis [5]. To describe complicated problems, interval-valued fuzzy numbers were also studied in fuzzy risk analysis [4, 6]. Intuitionistic fuzzy set

X. Zhou (✉) · S. Wang · C. Zhang
College of Information Engineering, Dalian University, Dalian 116622, China
e-mail: zhouxiaojie@dlu.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_11

is an extension of fuzzy set, by introducing nonmembership degree [7]. Intuitionistic fuzzy sets have been widely used in decision and management. In this paper, a new fuzzy risk analysis method is presented based on trapezoidal intuitionistic fuzzy numbers.

2 Preliminaries

The intuitionistic fuzzy number is an important concept of the intuitionistic fuzzy modeling. A trapezoidal intuitionistic fuzzy number $\tilde{A} = \langle (a_1, a_2, a_3, a_4); w_{\tilde{A}}, u_{\tilde{A}} \rangle$ is a special intuitionistic fuzzy set on the real number set $\mathbf{R}$, whose membership and nonmembership functions are defined as follows [8]:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < a_1 \\ (x - a_1) w_{\tilde{A}} / (a_2 - a_1), & a_1 \le x < a_2 \\ w_{\tilde{A}}, & a_2 \le x < a_3 \\ (a_4 - x) w_{\tilde{A}} / (a_4 - a_3), & a_3 \le x \le a_4 \\ 0, & x > a_4 \end{cases} \tag{1}$$

$$v_{\tilde{A}}(x) = \begin{cases} 1, & x < a_1 \\ [a_2 - x + u_{\tilde{A}}(x - a_1)]/(a_2 - a_1), & a_1 \le x < a_2 \\ u_{\tilde{A}}, & a_2 \le x < a_3 \\ [x - a_3 + u_{\tilde{A}}(a_4 - x)]/(a_4 - a_3), & a_3 \le x \le a_4 \\ 1, & x > a_4 \end{cases} \tag{2}$$

Assume that there are two trapezoidal intuitionistic fuzzy numbers, $\tilde{A} = \langle (a_1, a_2, a_3, a_4); w_{\tilde{A}}, u_{\tilde{A}} \rangle$ and $\tilde{B} = \langle (b_1, b_2, b_3, b_4); w_{\tilde{B}}, u_{\tilde{B}} \rangle$, the arithmetic operations between $\tilde{A}$ and $\tilde{B}$ are as follows:

$$\tilde{A} + \tilde{B} = \langle (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4); w_{\tilde{A}} \wedge w_{\tilde{B}}, u_{\tilde{A}} \vee u_{\tilde{B}} \rangle, \tag{3}$$

$$\tilde{A} - \tilde{B} = \langle (a_1 - b_1, a_2 - b_2, a_3 - b_3, a_4 - b_4); w_{\tilde{A}} \wedge w_{\tilde{B}}, u_{\tilde{A}} \vee u_{\tilde{B}} \rangle. \tag{4}$$

Suppose $a_1 > 0$ and $b_1 > 0$, then

$$\tilde{A} \times \tilde{B} = \langle (a_1 b_1, a_2 b_2, a_3 b_3, a_4 b_4); w_{\tilde{A}} \wedge w_{\tilde{B}}, u_{\tilde{A}} \vee u_{\tilde{B}} \rangle, \tag{5}$$

$$\tilde{A} / \tilde{B} = \langle (a_1 / b_4, a_2 / b_3, a_3 / b_2, a_4 / b_1); w_{\tilde{A}} \wedge w_{\tilde{B}}, u_{\tilde{A}} \vee u_{\tilde{B}} \rangle. \tag{6}$$

3 Weighted Similarity Measure Between Trapezoidal Intuitionistic Fuzzy Numbers

The center-of-gravity method is very useful to deal with defuzzification problems and fuzzy ranking problems. For a generalized trapezoidal fuzzy number $\hat{A} = (a_1, a_2, a_3, a_4; w_{\hat{A}})$, with membership function $\mu_{\hat{A}}(x)$, the center-of-gravity of $\hat{A}$ is [1, 9]

$$COG(\hat{A}) = (x^*_{\hat{A}}, y^*_{\hat{A}}) \tag{7}$$

where

$$x^*_{\hat{A}} = [y^*_{\hat{A}}(a_3 + a_2) + (a_4 + a_1)(w_{\hat{A}} - y^*_{\hat{A}})]/(2w_{\hat{A}}),$$
$$y^*_{\hat{A}} = \begin{cases} w_{\hat{A}}\{(a_3 - a_2)/[(a_3 - a_2) + (a_4 - a_1)] + 1\}/3, & a_1 \neq a_4 \\ w_{\hat{A}}/2, & a_1 = a_4 \end{cases}.$$

In [1], Chen and Chen calculate the similarity degree of two generalized trapezoidal $\hat{A} = (a_1, a_2, a_3, a_4; w_{\hat{A}})$ and $\hat{B} = (b_1, b_2, b_3, b_4; w_{\hat{B}})$ as

$$S(\hat{A}, \hat{B}) = \left[1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4}\right] \cdot (1 - |x^*_{\hat{A}} - x^*_{\hat{B}}|)^{B(S_{\hat{A}}, S_{\hat{B}})} \cdot \frac{\min(y^*_{\hat{A}}, y^*_{\hat{B}})}{\max(y^*_{\hat{A}}, y^*_{\hat{B}})}, \tag{8}$$

where $(x^*_{\hat{A}}, y^*_{\hat{A}})$ and $(x^*_{\hat{B}}, y^*_{\hat{B}})$ are the centers of gravity of $\mu_{\hat{A}}(x)$ and $\mu_{\hat{B}}(x)$ respectively. $B(S_{\hat{A}}, S_{\hat{B}})$ is defined as follows:

$$B(S_{\hat{A}}, S_{\hat{B}}) = \begin{cases} 1, & S_{\hat{A}} + S_{\hat{B}} > 0 \\ 0, & S_{\hat{A}} + S_{\hat{B}} = 0 \end{cases}, \tag{9}$$

where $S_{\hat{A}}$ and $S_{\hat{B}}$ are the lengths of the bases of the generalized trapezoidal fuzzy numbers $\hat{A}$ and $\hat{B}$, respectively, defined as $S_{\hat{A}} = a_4 - a_1$ and $S_{\hat{B}} = b_4 - b_1$.

Inspired by the above similarity of generalized trapezoidal fuzzy numbers, we define the similarity of trapezoidal intuitionistic fuzzy numbers. Assume that there are two trapezoidal intuitionistic fuzzy numbers, $\tilde{A} = \langle(a_1, a_2, a_3, a_4); w_{\tilde{A}}, u_{\tilde{A}}\rangle$ and $\tilde{B} = \langle(b_1, b_2, b_3, b_4); w_{\tilde{B}}, u_{\tilde{B}}\rangle$, $\mu_{\tilde{A}}(x)$ and $v_{\tilde{A}}(x)$ are the membership and nonmembership functions of $\tilde{A}$ respectively, and $\mu_{\tilde{B}}(x)$ and $v_{\tilde{B}}(x)$ for $\tilde{B}$. $(x^*_{\tilde{A},\mu}, y^*_{\tilde{A},\mu})$ and $(x^*_{\tilde{B},\mu}, y^*_{\tilde{B},\mu})$ are the centers of gravity of $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$, and $(x^*_{\tilde{A},v}, y^*_{\tilde{A},v})$ and $(x^*_{\tilde{B},v}, y^*_{\tilde{B},v})$ are the centers of gravity of $v_{\tilde{A}}(x)$ and $v_{\tilde{B}}(x)$. According to Eq. (8), we can obtain the similarity of $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ as:

$$S_\mu(\tilde{A}, \tilde{B}) = \left[1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4}\right] \cdot (1 - |x^*_{\tilde{A},\mu} - x^*_{\tilde{B},\mu}|)^{B(S_{\tilde{A}}, S_{\tilde{B}})} \cdot \frac{\min(y^*_{\tilde{A},\mu}, y^*_{\tilde{B},\mu})}{\max(y^*_{\tilde{A},\mu}, y^*_{\tilde{B},\mu})}, \tag{10}$$

And the similarity of $v_{\tilde{A}}(x)$ and $v_{\tilde{B}}(x)$ as:

$$S_v(\tilde{A},\tilde{B}) = \left[1 - \frac{\sum_{i=1}^{4}|a_i - b_i|}{4}\right] \cdot (1 - |x^*_{\tilde{A},v} - x^*_{\tilde{B},v}|)^{B(S_{\tilde{A}},S_{\tilde{B}})} \cdot \frac{\min(y^*_{\tilde{A},v}, y^*_{\tilde{B},v})}{\max(y^*_{\tilde{A},v}, y^*_{\tilde{B},v})}. \quad (11)$$

Then the similarity of trapezoidal intuitionistic fuzzy numbers $\tilde{A}$ and $\tilde{B}$ can be defined as:

$$S(\tilde{A},\tilde{B}) = \alpha S_\mu(\tilde{A},\tilde{B}) + (1-\alpha)S_v(\tilde{A},\tilde{B}), \quad (12)$$

where α is the weight coefficient of membership function with $0 \le \alpha \le 1$. For a trapezoidal intuitionistic fuzzy number $\tilde{A}$, it is obvious that $x^*_{\tilde{A},\mu} = x^*_{\tilde{A},v}$, named as $x^*_{\tilde{A}}$, then Eq. (12) can be converted into:

$$S(\tilde{A},\tilde{B}) = \left[1 - \frac{\sum_{i=1}^{4}|a_i - b_i|}{4}\right] \cdot (1 - |x^*_{\tilde{A}} - x^*_{\tilde{B}}|)^{B(S_{\tilde{A}},S_{\tilde{B}})} \cdot \left(\alpha \frac{\min(y^*_{\tilde{A},\mu}, y^*_{\tilde{B},\mu})}{\max(y^*_{\tilde{A},\mu}, y^*_{\tilde{B},\mu})} + (1-\alpha)\frac{\min(y^*_{\tilde{A},v}, y^*_{\tilde{B},v})}{\max(y^*_{\tilde{A},v}, y^*_{\tilde{B},v})}\right). \quad (13)$$

In the following, we give an example to illustrate the similarity measure between two trapezoidal intuitionistic fuzzy numbers.

Example 1 Assume that there are two trapezoidal intuitionistic fuzzy numbers $\tilde{A}$ and $\tilde{B}$, where

$$\tilde{A} = \langle(1,2,3,4); 0.45, 0.55\rangle, \quad (14)$$

$$\tilde{B} = \langle(1,1.5,2.5,4); 0.2, 0.5\rangle, \quad (15)$$

as shown in Fig. 1a. According to Eq. (7), we obtain that $x^*_{\tilde{A}} = 0.25$, $x^*_{\tilde{B}} = 0.2917$, $y^*_{\tilde{A},\mu} = 0.1875$, $y^*_{\tilde{A},v} = 0.8125$, $y^*_{\tilde{B},\mu} = 0.0822$, $y^*_{\tilde{B},v} = 0.7917$. The similarity degree

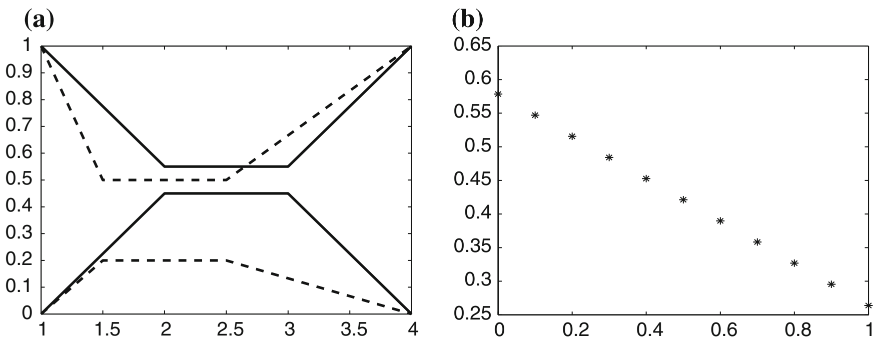


Fig. 1 Degree of similarity between $\tilde{A}$ and $\tilde{B}$. **a** Trapezoidal intuitionistic fuzzy numbers $\tilde{A}$ and $\tilde{B}$, *solid line* $\tilde{A}$, *dashed line* $\tilde{B}$; **b** Degree of similarity between $\tilde{A}$ and $\tilde{B}$ with regard to α

with regard to different weight coefficient α is shown in Fig. 1b. In this example, since the similarity degree of nonmemberships between $\tilde{A}$ and $\tilde{B}$ is more similar than the one of memberships, the similarity degree between $\tilde{A}$ and $\tilde{B}$ decreases with the increasing of α.

4 Fuzzy Risk Analysis Based on the Similarity Measure Between Intuitionistic Fuzzy Numbers

In this section, we will apply the proposed similarity measure between intuitionistic fuzzy numbers to fuzzy risk analysis. Since an intuitionistic fuzzy number has membership and nonmembership functions, which makes it more suitable for representing the opinion of decision maker than generalized trapezoidal fuzzy number. Assume that there is a component A consisting of n sub-components $A_1, A_2, \ldots, A_n$. $\tilde{R}_i$ and $\tilde{W}_i$ denote the probability of failure and the severity of loss of the sub-component $A_i, i \leq i \leq n$. The total risk of failure can be computed by:

$$\tilde{R} = \frac{\sum_{i=1}^{n} \tilde{W}_i \times \tilde{R}_i}{\sum_{i=1}^{n} \tilde{W}_i} \tag{16}$$

Then the total risk of failure can be translated to linguistic term by similarity measure [2].

Example 2 This example is referred to [5], the structure of the fuzzy risk analysis is shown as Fig. 2. Assume that the component A consists of three sub-components A_1, A_2, A_3. $\tilde{R}_i$ and $\tilde{W}_i$ $(i = 1, 2, 3)$ are illustrated in Table 1.

In order to apply the trapezoidal intuitionistic fuzzy numbers to fuzzy risk analysis, the linguistic values are represented by trapezoidal intuitionistic fuzzy numbers as shown in Table 2.

According to the arithmetic operations of trapezoidal intuitionistic fuzzy numbers and Eq. (14), we can obtain $\tilde{R} = \langle (0.1917, 0.3359, 0.5836, 1.0385); 0.5, 0.5 \rangle$. Let $\alpha = 0.5$, then the similarity degree of each linguistic term can be calculated by Eq. (13), i.e.,

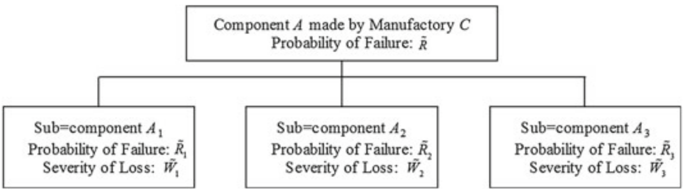


Fig. 2 Structure of fuzzy risk analysis [2]

Table 1 Linguistic values of $\tilde{R}_i$ and $\tilde{W}_i$

Sub-component A_i	$\tilde{W}_i$	$\tilde{R}_i$
A_1	Low	Low
A_2	Fairly-high	Medium
A_3	Very-low	High

Table 2 A 9-member linguistic term set

Linguistic terms	Trapezoidal intuitionistic fuzzy numbers
Absolutely-low	$\langle(0, 0, 0, 0); 1, 0\rangle$
Very-low	$\langle(0, 0, 0.075, 0.15); 0.9, 0.05\rangle$
Low	$\langle(0.1, 0.175, 0.225, 0.3); 0.8, 0.15\rangle$
Fairly-low	$\langle(0.25, 0.325, 0.375, 0.45); 0.65, 0.25\rangle$
Medium	$\langle(0.4, 0.475, 0.525, 0.6); 0.45, 0.45\rangle$
Fairly-high	$\langle(0.55, 0.625, 0.675, 0.75); 0.65, 0.25\rangle$
High	$\langle(0.7, 0.775, 0.825, 0.9); 0.8, 0.15\rangle$
Very-high	$\langle(0.85, 0.925, 1, 1); 0.9, 0.05\rangle$
Absolutely-high	$\langle(1, 1, 1, 1); 1, 0\rangle$

$$
\begin{aligned}
&S(\tilde{R}, \text{Absolutely}-\text{low}) = 0.1076,\\
&S(\tilde{R}, \text{Very}-\text{low}) = 0.2494,\\
&S(\tilde{R}, \text{Low}) = 0.4239,\\
&S(\tilde{R}, \text{Fairly}-\text{low}) = 0.6171,\\
&S(\tilde{R}, \text{Medium}) = 0.7381,\\
&S(\tilde{R}, \text{Fairly}-\text{high}) = 0.6610,\\
&S(\tilde{R}, \text{High}) = 0.4954,\\
&S(\tilde{R}, \text{Very}-\text{high}) = 0.3324,\\
&S(\tilde{R}, \text{Avsolutely}-\text{high}) = 0.1482.
\end{aligned}
$$

Since $S(\tilde{R}, \text{Medium}) = 0.7381$ has the largest value, $\tilde{R}$ is translated into the linguistic term "Medium". It means that the probability of failure of the component A is medium. This result coincides with the one shown in [5].

In reality, $\tilde{W}_i$ and $\tilde{R}_i$ may be represented as trapezoidal intuitionistic fuzzy numbers rather than linguistic values. In example 2, let $\tilde{R}_1 = \langle(0.3, 0.4, 0.5, 0.6); 0.4, 0.4\rangle$, $\tilde{R}_2 = \langle(0.4, 0.5, 0.55, 0.6); 0.6, 0.2\rangle$, $\tilde{R}_3 = \langle(0.6, 0.7, 0.75, 0.9); 0.4, 0.6\rangle$, and $\tilde{W}_1 = \langle(0, 0.05, 0.1, 0.2); 1, 0\rangle$, $\tilde{W}_2 = \langle(0.4, 0.5, 0.6, 0.7); 0.5, 0.2\rangle$, $\tilde{W}_3 = \langle(0.5, 0.6, 0.7, 0.75); 0.3, 0.5\rangle$. Then we can also calculate the degree of similarity of $\tilde{R}$ to every linguistic value shown as Fig. 3. From Fig. 3, we can see that if $\alpha = 0.3$, $S(\tilde{R}, \text{Fairly}-\text{high})$ is the largest similarity, while if $\alpha = 0.5$ or 0.7, $S(\tilde{R}, \text{Medium})$ is the largest similarity. The coefficient α reflects the weight of membership function relative to nonmembership function, and may affect the result of risk analysis.

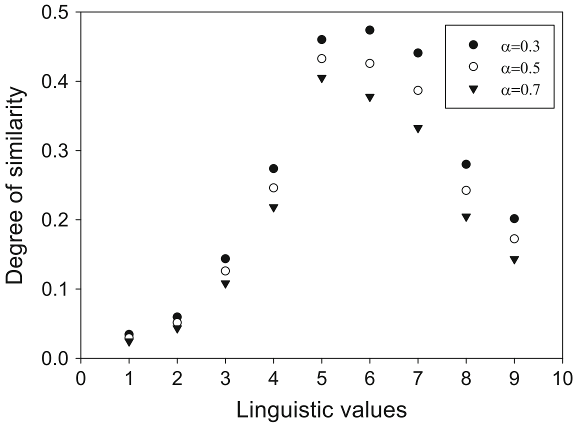


Fig. 3 Degree of similarity of $\tilde{R}$ to every linguistic value with respect to α

5 Conclusion

In this paper, a new fuzzy risk analysis method is proposed based on trapezoidal intuitionistic fuzzy numbers. A weighted similarity measure between trapezoidal intuitionistic fuzzy numbers is presented based on centers-of-gravity of membership and nonmembership functions. The new similarity measure is applied to fuzzy risk analysis based on trapezoidal intuitionistic fuzzy numbers. The weighted similarity improves the flexibility of a system. Further work may focus on the determination of the weight coefficient.

References

1. Chen, S.J., Chen, S.M.: Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. IEEE Trans. Fuzzy Syst. **11**(1), 45–56 (2003)
2. Schmucker, K.J.: Fuzzy Sets, Natural Language Computations and Risk, Analysis. Computer Science Press, MD (1984)
3. Xu, Z.Y., Shang, S.C., Qian, W.B., Shu, W.H.: A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers. Expert Syst. Appl. **37**(3), 1920–1927 (2010)
4. Chen, S.M., Sanguansat, K.: Analyzing fuzzy risk based on similarity measures between interval-valued fuzzy numbers. Expert Syst. Appl. **38**(7), 8612–8621 (2011)
5. Wei, S.H., Chen, S.M.: A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. Expert Syst. Appl. **36**(1), 589–598 (2009)
6. Chen, S.M., Chen, J.H.: Fuzzy risk analysis based on similarity measures between interval-valued fuzzy numbers and interval-valued fuzzy number arithmetic operators. Expert Syst. Appl. **36**(3), 6309–6317 (2009)
7. Atnassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
8. Li, D.F.: Decision and Game Theory in Management with Intuitionistic Fuzzy Sets. Springer, Berlin (2014)
9. Zhang, C., Li, T., Liu, X.Z.: Fuzzy risk analysis based on COG method. J. Liaoning Tech. Univ. (Nat. Sci.) **31**(2), 240–243 (2012)

On Fuzzy Soft Relation

Yu-hong Zhang, Xue-hai Yuan and Zun-quan Xia

Abstract Soft set theory, proposed by Molodtsov, has been regarded as a general mathematical tool for dealing with uncertainties. Research on soft set has received much attention in recent years. In this paper, we first give the notions of soft relation and fuzzy soft relation. Furthermore, the concepts of projection and section of soft relation are introduced and in the meantime, some of their properties are studied. Moreover, we define fuzzy soft linear transformation and get some conclusions.

Keywords Soft set · Soft relation · Fuzzy soft relation · Fuzzy soft linear transformation

1 Introduction

Molodtsov [1] initiated the concept of soft set theory as a completely new approach for modeling uncertainties in 1999. After Molodtsov's work, Maji et al. [2] introduced several operations on soft sets. Since then, some authors [3, 4] introduced and studied soft set operations as well. Maji et al. [5] also presented the definition of fuzzy soft set which many researchers [6–12] worked on presently. Aktas and Cagman [13] gave a definition of soft group. Some scholars [14–22] also discussed the algebraic structure of soft set. The application of soft set theory was also extended to incomplete information system [23] and decision making problems [24–26].

Y. Zhang (✉) · Z. Xia
School of Mathematical Sciences, Dalian University of Technology,
Dalian 116024, People's Republic of China
e-mail: yuhz@dlut.edu.cn

Y. Zhang
Department of Basic Education, City Institute, Dalian University of Technology,
Dalian 116600, People's Republic of China

X. Yuan
Faculty of Electronic Information and Electrical Engineering,
Dalian University of Technology, Dalian 116024, People's Republic of China
e-mail: yuanxh@dlut.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_12

Majumdar and Samanta [27] defined soft mapping in 2010. The main purpose of this paper is to introduce the concept of fuzzy soft relation, which is based on the direct relationship between mapping and relation.

The paper is organized as follows. Section 2 introduces some preliminary results. Section 3 proposes the definitions of soft relation and fuzzy soft relation and studies some related properties. Finally Sect. 4 concludes from this work.

2 Preliminaries

Definition 1 [1]. *Let U be an initial universe set and E be a set of parameters. Let $\mathcal{P}(U)$ denote the power set of U and $A \subset E$. A pair (F, A) is called a soft set over U iff F is a mapping given by $F : A \rightarrow \mathcal{P}(U)$.*

Example 2 Suppose a soft set (F, E) describes the attractiveness of the shirts which the authors are going to wear.

U = the set of all shirts under consideration $= \{x_1, x_2, x_3, x_4, x_5\}$

$E = \{$colorful,bright,cheap,warm$\} = \{e_1, e_2, e_3, e_4\}$

Let $F(e_1) = \{x_1, x_2\}, F(e_2) = \{x_1, x_2, x_3\}, F(e_3) = \{x_4\}, F(e_4) = \{x_2, x_5\}$
So,the soft set (F, E) is a family $\{F(e_i), i = 1, 2, 3, 4\}$ of $\mathcal{P}(U)$.

Definition 3 [27]. *Let A, B be two non-empty set and E be a parameter set. Then the mapping $F : E \rightarrow \mathcal{P}(B^A)$ is called a soft mapping from A to B under E, where B^A is the collection of all mappings from A to B.*

Actually a soft mapping F from A to B under E is a soft set over B^A.

Definition 4 [28]. *Let (F, A) and (G, B) be two soft sets over U, then a relation from (F, A) to (G, B) is a soft subset of $(F, A) \times (G, B)$.*

Definition 5 [29]. *A fuzzy soft relation is defined as soft set over the fuzzy power set of the cartesian product of two crisp sets. Let X and Y be two crisp sets and E is the set of parameters, then a function $R : E \rightarrow I^{X \times Y}$ is called a fuzzy soft relation.*

The next definitions and results are from [30].

Let $\mathcal{F}(X \times Y)$ denote the set of fuzzy subsets over $X \times Y, R \in \mathcal{F}(X \times Y)$,
$x \in X, y \in Y$, we set
$(R \mid_x)(y) = R(x, y), (R \mid_y)(x) = R(x, y)$.
$(R \mid_{[x]})(y) = 1 - R(x, y), (R \mid_{[y]})(x) = 1 - R(x, y)$.
$(R_X)(x) = \bigvee_{y \in Y} R(x, y), (R_Y)(y) = \bigvee_{x \in X} R(x, y)$.
$(R_{[X]})(x) = \bigwedge_{y \in Y} R(x, y), (R_{[Y]})(y) = \bigwedge_{x \in X} R(x, y)$.

Property 6 $R_2 \subseteq R_1 \Rightarrow (R_2)_X \subseteq (R_1)_X, (R_2)_Y \subseteq (R_1)_Y, (R_2)_x \subseteq (R_1)_x$,
$(R_2)_y \subseteq (R_1)_y, (R_2)_{[x]} \supseteq (R_1)_{[x]}, (R_2)_{[y]} \supseteq (R_1)_{[y]}$.

Property 7 *(1)* $R_{[X]} \subseteq R_X$.
(2) $R\mid_{[x]} = R^c\mid_x, R^c\mid_{[x]} = (R\mid_{[x]})^c, R^c\mid_x = (R\mid_x)^c$.
(3) $R_{[X]} = ((R^c)_X)^c$.
(4) $(R \cup S)_X = R_X \cup S_X, (R \cap S)_X \subseteq R_X \cap S_X$
$(\bigcup_{t\in T} R_t)_X = \bigcup_{t\in T}(R_t)_X, (\bigcap_{t\in T} R_t)_X \subseteq \bigcap_{t\in T}(R_t)_X$.
(5) $(R \cup S)_{[X]} \supseteq R_{[X]} \cup S_{[X]}, (R \cap S)_{[X]} = R_{[X]} \cap S_{[X]}$.
$(\bigcup_{t\in T} R_t)_{[X]} \supseteq \bigcup_{t\in T}(R_t)_{[X]}, (\bigcap_{t\in T} R_t)_{[X]} = \bigcap_{t\in T}(R_t)_{[X]}$.
(6) $(\bigcup_{t\in T} R_t)\mid_x = \bigcup_{t\in T}(R_t)\mid_x, (\bigcap_{t\in T} R_t)\mid_x = \bigcap_{t\in T}(R_t)\mid_x$.
(7) $(\bigcup_{t\in T} R_t)\mid_{[x]} = \bigcap_{t\in T}(R_t)\mid_{[x]}, (\bigcap_{t\in T} R_t)\mid_{[x]} = \bigcup_{t\in T}(R_t)\mid_{[x]}$.

Let R, S, T and Q be fuzzy relations, then we have that

(1) $(R \circ S) \circ T = R \circ (S \circ T)$.
(2) $(R \cup S) \circ T = (R \circ T) \cup (S \circ T), T \circ (R \cup S) = (T \circ R) \cup (T \circ S)$.
$(\bigcup_{t\in T} R_t) \circ T = \bigcup_{t\in T}(R_t \circ T), S \circ (\bigcup_{t\in T} R_t) = \bigcup_{t\in T}(S \circ R_t)$.
(3) $(\lambda R) \circ S = \lambda(R \circ S) = R \circ (\lambda S)$ $(\lambda \in [0, 1])$
$(\bigcup_{r\in\Gamma} \lambda_r R^{(r)}) \circ T = \bigcup_{r\in\Gamma} \lambda_r (R^{(r)} \circ T), S \circ (\bigcup_{r\in\Gamma} \lambda_r R^{(r)}) = \bigcup_{r\in\Gamma} \lambda_r (S \circ R^{(r)})$.
(4) $Q \subseteq R \Rightarrow Q \circ T \subseteq R \circ T, S \circ Q \subseteq S \circ R, Q^n \subseteq R^n$,
$(\bigcap_{r\in\Gamma} R^{(r)}) \circ T \subseteq \bigcap_{r\in\Gamma}(R^{(r)} \circ T)$.

3 Soft Relation and Fuzzy Soft Relation

In this section we introduce the notions of soft relation and fuzzy soft relation and study their properties. Let X be the universal set and E be a parameter set. Then the pair (X, E) will be called a soft universe. Throughout this section we assume that (X, E) is our soft universe.

Definition 8 Let A, B be two non-empty set and E be a parameter set. Then the mapping

$$F : E \to \mathcal{P}([0, 1]^{A\times B})$$

is called a soft relation from A to B under E, where $[0, 1]^{A\times B}$ is the collection of all fuzzy relations from A to B.
Note Actually a soft relation F from A to B under E is a soft set over $[0, 1]^{A\times B}$, which is different from the definition of soft set relation in [28]. In fact, let (F, A) and (G, B) be two soft sets over U, then a soft set relation from (F, A) to (G, B) is a soft subset of $(F, A) \times (G, B)$, where $(F, A) \times (G, B) = (H, A \times B)$, where $H : A \times B \to \mathcal{P}(U \times U)$ and $H(a, b) = F(a) \times G(b)$, where $(a, b) \in A \times B$.

Example 9 Let $F : E \to \mathcal{P}(B^A)$ is a soft mapping from A to B under E. Then soft mapping is a special case of soft relation.

Example 10 Let $E = \{e_1, e_2\}, A = \{a_1, a_2\}, B = \{b_1, b_2, b_3\}$.

Let $R_1, R_2, R_3, R_4 \in [0,1]^{A\times B}$ be defined as follows:

$$R_1 = \begin{pmatrix} 1 & 0.1 & 0.2 \\ 0 & 1 & 0.3 \end{pmatrix}, R_2 = \begin{pmatrix} 0 & 1 & 0.3 \\ 1 & 0.1 & 0.2 \end{pmatrix},$$

$$R_3 = \begin{pmatrix} 0.1 & 1 & 0.2 \\ 0.3 & 0 & 1 \end{pmatrix}, R_4 = \begin{pmatrix} 0 & 0.5 & 1 \\ 0.6 & 1 & 0.7 \end{pmatrix},$$

Let $F : E \to \mathcal{P}([0,1]^{A\times B})$ be defined as follows:
$F(e_1) = \{R_1, R_4\}, F(e_2) = \{R_1, R_2, R_3\}$.
Then F is a soft relation from A to B under E.

Definition 11 Let A, B be two non-empty set and E be a parameter set. Then the mapping

$$F : E \to \mathcal{F}([0,1]^{A\times B})$$

is called a fuzzy soft relation from A to B under E.
Note This is different from the definition of fuzzy soft relation in [29], which is defined as soft set over the fuzzy power set of the cartesian product of two crisp sets.

Definition 12 Let F be a soft relation from A to A under E, we say that

(1) F is reflexive$\Leftrightarrow \forall R \in F(e)$ is reflexive, $\forall e \in E$
(2) F is symmetric$\Leftrightarrow \forall R \in F(e)$ is symmetric, $\forall e \in E$
(3) F is transitive$\Leftrightarrow \forall R \in F(e)$ is transitive, $\forall e \in E$

A soft relation F is called a soft equivalence relation from A to A under E if it is reflexive, symmetric and transitive.

Definition 13 Let F_1 be a soft relation from A to B under E, F_2 be a soft relation from B to C under E. Then a new soft relation, the composition of F_1 and F_2 expressed as $F_1 \circ F_2$ from A to C under E is defined as follows:

$$(F_1 \circ F_2)(e) = \{R_1 \circ R_2 \mid R_1 \in F_1(e), R_2 \in F_2(e)\}$$

Let F and G be soft relations from A to B under E, we set
$F \le G \Leftrightarrow \forall e \in E, F(e) \subseteq G(e); F = G \Leftrightarrow \forall e \in E, F(e) = G(e);$
$F^c(e) = \{R^c \mid R \in F(e)\}; (F \cup G)(e) = F(e) \cup G(e);$
$(F \cap G)(e) = F(e) \cap G(e); (\lambda F)(e) = \{\lambda R \mid R \in F(e)\}.$

Property 14 *Let F, G and S be soft relations,*

(1) $(F \circ G) \circ S = F \circ (G \circ S)$.
(2) $(F \cup G) \circ S = (F \circ S) \cup (G \circ S), S \circ (F \cup G) = (S \circ F) \cup (S \circ G)$.
$(\bigcup_{t\in T} F_t) \circ G = \bigcup_{t\in T}(F_t \circ G), S \circ (\bigcup_{t\in T} F_t) = \bigcup_{t\in T}(S \circ F_t)$.

(3) $(\lambda F)\circ G = \lambda(F\circ G) = F\circ(\lambda G)(\lambda \in [0,1])$
$(\bigcup_{r\in\Gamma} \lambda_r F^{(r)})\circ G = \bigcup_{r\in\Gamma} \lambda_r(F^{(r)}\circ G), S\circ(\bigcup_{r\in\Gamma} \lambda_r F^{(r)}) = \bigcup_{r\in\Gamma} \lambda_r(S\circ F^{(r)})$.
(4) $F_1 \le F_2 \Rightarrow F_1\circ G \le F_2\circ G, S\circ F_1 \le S\circ F_2, F_1^n \le F_2^n$,

$(\bigcap_{r\in\Gamma} F^{(r)})\circ G \le \bigcap_{r\in\Gamma}(F^{(r)}\circ G)$.

Definition 15 Let F be a soft relation from A to B under E. For $a \in A$ and $b \in B$, we set

$$
\begin{aligned}
F\,|_a &: E \to \mathcal{P}([0,1]^B)\\
&\quad e \mapsto F\,|_a\,(e) = \{R\,|_a \mid R \in F(e)\}\\
F\,|_b &: E \to \mathcal{P}([0,1]^A)\\
&\quad e \mapsto F\,|_b\,(e) = \{R\,|_b \mid R \in F(e)\}\\
F\,|_{[a]} &: E \to \mathcal{P}([0,1]^B)\\
&\quad e \mapsto F\,|_{[a]}\,(e) = \{R\,|_{[a]} \mid R \in F(e)\}\\
F\,|_{[b]} &: E \to \mathcal{P}([0,1]^A)\\
&\quad e \mapsto F\,|_{[b]}\,(e) = \{R\,|_{[b]} \mid R \in F(e)\}\\
F_A &: E \to \mathcal{P}([0,1]^A)\\
&\quad e \mapsto F_A(e) = \{R_A \mid R \in F(e)\}\\
F_B &: E \to \mathcal{P}([0,1]^B)\\
&\quad e \mapsto F_B(e) = \{R_B \mid R \in F(e)\}\\
F_{[A]} &: E \to \mathcal{P}([0,1]^A)\\
&\quad e \mapsto F_{[A]}(e) = \{R_{[A]} \mid R \in F(e)\}\\
F_{[B]} &: E \to \mathcal{P}([0,1]^B)\\
&\quad e \mapsto F_{[B]}(e) = \{R_{[B]} \mid R \in F(e)\}
\end{aligned}
$$

Property 16 *Let F_1, F_2 be soft relations from A to B under E, then we have that*
$F_2 \le F_1 \Rightarrow (F_2)_A \le (F_1)_A, (F_2)_B \le (F_1)_B, F_2\,|_a \le F_1\,|_a, F_2\,|_b \le F_1\,|_b$, $F_2\,|_{[a]} \le F_1\,|_{[a]}, F_2\,|_{[b]} \le F_1\,|_{[b]}$.
i.e. $F_2(e) \subseteq F_1(e) \Rightarrow (F_2)_A(e) \subseteq (F_1)_A(e), (F_2)_B(e) \subseteq (F_1)_B(e)$, $F_2\,|_a\,(e) \subseteq F_1\,|_a\,(e), F_2\,|_b\,(e) \subseteq F_1\,|_b\,(e), F_2\,|_{[a]}\,(e) \subseteq F_1\,|_{[a]}\,(e)$, $F_2\,|_{[b]}\,(e) \subseteq F_1\,|_{[b]}\,(e)$.

Property 17 *Let F be a soft relation from A to B under E, then we have that*

(1) $F\,|_{[a]} = F^c\,|_a, F^c\,|_{[a]} = (F\,|_{[a]})^c, F^c\,|_a = (F\,|_a)^c$.
(2) $F_{[A]} = ((F^c)_A)^c$.
(3) $(F\cup G)_A = F_A \cup G_A, (F\cap G)_A = F_A \cap G_A$.
$(\bigcup_{t\in T} F_t)_A = \bigcup_{t\in T}(F_t)_A, (\bigcap_{t\in T} F_t)_A = \bigcap_{t\in T}(F_t)_A$.

(4) $(F \cup G)_{[A]} = F_{[A]} \cup G_{[A]}, (F \cap G)_{[A]} = F_{[A]} \cap G_{[A]}$.
$(\bigcup_{t\in T} F_t)_{[A]} = \bigcup_{t\in T}(F_t)_{[A]}, (\bigcap_{t\in T} F_t)_{[A]} = \bigcap_{t\in T}(F_t)_{[A]}$.
(5) $(\bigcup_{t\in T} F_t) \mid_a = \bigcup_{t\in T}(F_t) \mid_a, (\bigcap_{t\in T} F_t) \mid_a = \bigcap_{t\in T}(F_t) \mid_a$.
(6) $(\bigcup_{t\in T} F_t) \mid_{[a]} = \bigcap_{t\in T}(F_t) \mid_{[a]}, (\bigcap_{t\in T} F_t) \mid_{[a]} = \bigcup_{t\in T}(F_t) \mid_{[a]}$.

Definition 18 Let $F: E \to \mathcal{P}([0,1]^{A\times B})$ and $T: E \to \mathcal{P}([0,1]^A)$, we set

$$T\circ F : E \to \mathcal{P}([0,1]^B)$$
$$e \mapsto (T\circ F)(e)$$

where $(T\circ F)(e) = T(e)\circ F(e) = \{A\circ R \mid A \in T(e), R \in F(e)\}$
Let $T_1: E \to \mathcal{P}([0,1]^A)$ and $T_2: E \to \mathcal{P}([0,1]^A)$, we set

$$T_1 \cup T_2 : E \to \mathcal{P}([0,1]^A)$$
$$e \mapsto (T_1 \cup T_2)(e) = T_1(e) \cup T_2(e)$$

Let $T: E \to \mathcal{P}([0,1]^A)$, we set

$$\lambda T : E \to \mathcal{P}([0,1]^A)$$
$$e \mapsto (\lambda T)(e) = \{\lambda A \mid A \in T(e)\}$$

Let $F: E \to \mathcal{P}([0,1]^{A\times B})$ and $T: E \to \mathcal{P}([0,1]^A)$, we set

$$\lambda(T\circ F) : E \to \mathcal{P}([0,1]^B)$$
$$e \mapsto [\lambda(T\circ F)](e) = \{(\lambda A)\circ R \mid A \in T(e), R \in F(e)\}$$

Theorem 19 *Let* $T_1: E \to \mathcal{P}([0,1]^A)$*,* $T_2: E \to \mathcal{P}([0,1]^A)$ *and* $F: E \to \mathcal{P}([0,1]^{A\times B})$*, then*

(1) $(T_1 \cup T_2)\circ F = (T_1\circ F) \cup (T_2\circ F)$.
(2) $(\lambda T)\circ F = \lambda(T\circ F)$.
(3) $T_1 \leq T_2 \Rightarrow T_1\circ F \leq T_2\circ F$.

i.e. $T_1(e) \subseteq T_2(e) \Rightarrow (T_1\circ F)(e) \subseteq (T_2\circ F)(e)$.

Definition 20 Let $\mathcal{F}_A = \{T_A \mid T_A : E \to \mathcal{P}([0,1]^A)\}$, $\mathcal{F}_B = \{T_B \mid T_B : E \to \mathcal{P}([0,1]^B)\}$.
Let $\Gamma : \mathcal{F}_A \to \mathcal{F}_B$ is a mapping. If for $\lambda_t \in [0,1], {T_A}^{(t)} \in \mathcal{F}_A$ $(t \in T)$, we have that

$$\Gamma(\bigcup_{t\in T} \lambda_t {T_A}^{(t)}) = \bigcup_{t\in T} \lambda_t \Gamma({T_A}^{(t)})$$

then Γ is called a soft linear transformation from A to B.
Clearly, Γ has the following properties.

Property

(1) $\Gamma(T_A{}^{(1)} \cup T_A{}^{(2)}) = \Gamma(T_A{}^{(1)}) \cup \Gamma(T_A{}^{(2)})$.
(2) $\Gamma(\lambda\, T_A) = \lambda\, \Gamma(T_A)$.
(3) $T_A{}^{(1)} \leq T_A{}^{(2)} \Rightarrow \Gamma(T_A{}^{(1)}) \leq \Gamma(T_A{}^{(2)})$.

i.e. $T_A{}^{(1)}(e) \subseteq T_A{}^{(2)}(e) \Rightarrow \Gamma(T_A{}^{(1)})(e) \subseteq \Gamma(T_A{}^{(2)})(e)$.

Definition 21 Let $F: E \to \mathcal{F}([0,1]^{A\times B})$ and $T: E \to \mathcal{F}([0,1]^A)$, we set

$$T\circ F: E \to \mathcal{F}([0,1]^B)$$

$$(T\circ F)(e)(Y) = \bigvee_{Y=X\circ R} (T(e)(X) \wedge F(e)(R))$$

Let $T_1: E \to \mathcal{F}([0,1]^A)$ and $T_2: E \to \mathcal{F}([0,1]^A)$, we set

$$T_1 \cup T_2: E \to \mathcal{F}([0,1]^A)$$

$$(T_1 \cup T_2)(e)(X) = T_1(e)(X) \vee T_2(e)(X), \forall X \in [0,1]^A$$

Let $T: E \to \mathcal{F}([0,1]^A)$, we set

$$\lambda\, T: E \to \mathcal{F}([0,1]^A)$$
$$e \mapsto (\lambda\, T)(e) = \lambda\, T(e)$$

Let $F: E \to \mathcal{F}([0,1]^{A\times B})$ and $T : E \to \mathcal{F}([0,1]^A)$, we set

$$\lambda(T\circ F): E \to \mathcal{F}([0,1]^B)$$
$$e \mapsto [\lambda(T\circ F)](e) = \lambda(T\circ F)(e)$$

Theorem 22 *Let* $T_1: E \to \mathcal{F}([0,1]^A), T_2 : E \to \mathcal{F}([0,1]^A)$ *and* $F: E \to \mathcal{F}([0,1]^{A\times B})$*, then*

(1) $(T_1 \cup T_2)\circ F = (T_1\circ F) \cup (T_2\circ F)$.
(2) $(\lambda\, T)\circ F = \lambda(T\circ F)$.
(3) $T_1 \leq T_2 \Rightarrow T_1\circ F \leq T_2\circ F$.

i.e. $T_1(e)(X) \leq T_2(e)(X) \Rightarrow (T_1\circ F)(e)(Y) \leq (T_2\circ F)(e)(Y)$

Definition 23 Let $\mathcal{F}_A = \{T_A \mid T_A: E \to \mathcal{F}([0,1]^A)\}$, $\mathcal{F}_B = \{T_B \mid T_B: E \to \mathcal{F}([0,1]^B)\}$.
Let $\Gamma: \mathcal{F}_A \to \mathcal{F}_B$ is a mapping. If for $\lambda_t \in [0,1], T_A{}^{(t)} \in \mathcal{F}_A$ $(t \in T)$, we have that

$$\Gamma(\bigcup_{t\in T} \lambda_t T_A{}^{(t)}) = \bigcup_{t\in T} \lambda_t \Gamma(T_A{}^{(t)})$$

then Γ is called a fuzzy soft linear transformation from A to B.

Property *(1)* $\Gamma({T_A}^{(1)} \cup {T_A}^{(2)}) = \Gamma({T_A}^{(1)}) \cup \Gamma({T_A}^{(2)})$.
(2) $\Gamma(\lambda\, T_A) = \lambda\, \Gamma(T_A)$.
(3) ${T_A}^{(1)} \leq {T_A}^{(2)} \Rightarrow \Gamma({T_A}^{(1)}) \leq \Gamma({T_A}^{(2)})$.
i.e. ${T_A}^{(1)}(e)(X) \leq {T_A}^{(2)}(e)(X) \Rightarrow \Gamma({T_A}^{(1)})(e)(Y) \leq \Gamma({T_A}^{(2)})(e)(Y)$

Theorem 24 *Let $\Gamma : \mathcal{F}_A \to \mathcal{F}_B$ be a fuzzy soft linear transformation from A to B, then there is an unique fuzzy soft relation $F_\Gamma : E \to \mathcal{F}([0,1]^{A\times B})$ such that $\Gamma(T_A) = T_A \circ F_\Gamma$; On the contrary, let $F : E \to \mathcal{F}([0,1]^{A\times B})$ be a fuzzy soft relation, then there is an unique fuzzy soft linear transformation $\Gamma_F : \mathcal{F}_A \to \mathcal{F}_B$ such that $\Gamma_F(T_A) = T_A \circ F$.*

Proof We first show that $\{T_A(e) \mid T_A \in \mathcal{F}_A\} = \mathcal{F}([0,1]^A)$, for any $e \in E$.

In fact, for $H \in \mathcal{F}([0,1]^A)$, we set ${T_A}^H : E \to \mathcal{F}([0,1]^A)$ and ${T_A}^H(e) \equiv H$, $\forall e \in E$. Then $\mathcal{F}([0,1]^A) \supseteq \{T_A(e) \mid T_A \in \mathcal{F}_A\} \supseteq \{{T_A}^H(e) \mid H \in \mathcal{F}([0,1]^A)\} = \mathcal{F}([0,1]^A)$. It follows that $\{T_A(e) \mid T_A \in \mathcal{F}_A\} = \mathcal{F}([0,1]^A)$, $\forall e \in E$.

Let $f : \mathcal{F}([0,1]^A) \to \mathcal{F}([0,1]^B)$ be a mapping such that $f(T_A(e)) = \Gamma(T_A)(e)$, then

$$f(\bigcup_{t\in T} \lambda_t {T_A}^{(t)}(e)) = f((\bigcup_{t\in T} \lambda_t {T_A}^{(t)})(e)) = \Gamma(\bigcup_{t\in T} \lambda_t {T_A}^{(t)})(e)$$
$$= (\bigcup_{t\in T} \lambda_t \Gamma({T_A}^{(t)}))(e) = \bigcup_{t\in T} \lambda_t \Gamma({T_A}^{(t)})(e) = \bigcup_{t\in T} \lambda_t f({T_A}^{(t)}(e)).$$

Then f is a fuzzy soft linear transformation from A to B.

Let $F_\Gamma : E \to \mathcal{F}([0,1]^{A\times B})$ be a mapping and

$$F_\Gamma(e)(R) = \Gamma({T_A}^{\{X\}})(e)(Y) = f({T_A}^{\{X\}}(e))(Y) = f(\{X\})(Y).$$

For $T_A(e) \in \mathcal{F}([0,1]^A)$, let $\lambda_X = T_A(e)(X), \forall X \in [0,1]^A$

Then $(\bigcup_{X\in[0,1]^A} \lambda_X \{X\})(X') = \bigvee_{X\in[0,1]^A} (\lambda_X \wedge \{X\}(X')) = \lambda_{X'} \wedge \{X\}(X') = \lambda_{X'} \wedge 1 = \lambda_{X'}$
$= T_A(e)(X')$

Then $T_A(e) = \bigcup_{X\in[0,1]^A} \lambda_X \{X\}$

Then $\Gamma(T_A)(e) = f(T_A(e)) = f(\bigcup_{X\in[0,1]^A} \lambda_X \{X\})$.

Then $\Gamma(T_A)(e) = \bigcup_{X\in[0,1]^A} \lambda_X f(\{X\}) = \bigcup_{X\in[0,1]^A} \lambda_X f({T_A}^{\{X\}}(e))$.

Then $\Gamma(T_A)(e)(Y) = \bigvee_{X\in[0,1]^A} (\lambda_X \wedge f({T_A}^{\{X\}}(e))(Y))$
$= \bigvee_{Y=X\circ R} (T_A(e)(X) \wedge F_\Gamma(e)(R)) = (T_A \circ F_\Gamma)(e)(Y)$.

It follows that $\Gamma(T_A)(e) = T_A(e) \circ F_\Gamma(e)$ and consequently $\Gamma(T_A) = T_A \circ F_\Gamma$.

We need to prove that F_Γ is unique.

In fact, if the fuzzy soft relation $F : E \to \mathcal{F}([0,1]^{A\times B})$ satisfies $\Gamma(T_A) = T_A \circ F$, then

$$F_\Gamma(e)(R) = \Gamma({T_A}^{\{X\}})(e)(Y) = ({T_A}^{\{X\}}(e) \circ F(e))(Y)$$
$$= (\{X\} \circ F(e))(Y) = \bigvee_{Y=X'\circ R} (\{X\}(X') \wedge F(e)(R)) = F(e)(R)$$

Hence, $F(e) = F_\Gamma(e)$. It follows that $F = F_\Gamma$.

On the country, let $F : E \to \mathcal{F}([0,1]^{A\times B})$ be a fuzzy soft relation.

Let $\Gamma_F : \mathcal{F}_A \to \mathcal{F}_B$

$$T_A \mapsto \Gamma_F(T_A) = T_A \circ F$$

then Γ_F is a fuzzy soft linear transformation. In fact,

$$\Gamma(\bigcup_{t\in T} \lambda_t T_A{}^{(t)})(e) = (\bigcup_{t\in T} \lambda_t T_A{}^{(t)})(e) \circ F(e) = (\bigcup_{t\in T} \lambda_t T_A{}^{(t)}(e)) \circ F(e)$$
$$= \bigcup_{t\in T} (\lambda_t T_A{}^{(t)}(e)) \circ F(e) = \bigcup_{t\in T} \lambda_t (T_A{}^{(t)} \circ F)(e) = (\bigcup_{t\in T} \lambda_t (T_A{}^{(t)} \circ F))(e)$$
$$= (\bigcup_{t\in T} \lambda_t \Gamma(T_A{}^{(t)}))(e).$$

It follows that $\Gamma(\bigcup_{t\in T} \lambda_t T_A{}^{(t)}) = \bigcup_{t\in T} \lambda_t \Gamma(T_A{}^{(t)})$.
Since $\Gamma_F(T_A) = T_A \circ F$, Γ_F is unique.

Corollary. Let $\Gamma : \mathcal{F}_A \to \mathcal{F}_B$ be a soft linear transformation from A to B, then there is an unique soft relation $F_\Gamma : E \to \mathcal{P}([0, 1]^{A\times B})$ such that $\Gamma(T_A) = T_A \circ F_\Gamma$; On the contrary, let $F : E \to \mathcal{P}([0, 1]^{A\times B})$ be a soft relation, then there is an unique soft linear transformation $\Gamma_F : \mathcal{F}_A \to \mathcal{F}_B$ such that $\Gamma_F(T_A) = T_A \circ F$.

4 Conclusion

In this work, we introduced the concept of soft relation and fuzzy soft relation and studied some related properties. We also proposed the concept of fuzzy soft linear transformation and got some conclusions.

Acknowledgments The authors are grateful to the referees and the editor of this journal for their valuable comments and suggestions.

References

1. Molodtsov, D.: Soft set theory-first results. Comput. Math. Appl. **37**, 19–31 (1999)
2. Maji, P.K., Biswas, R., Roy, A.R.: Soft set theory. Comput. Math. Appl. **45**, 555–562 (2003)
3. Pei, D., Miao, D.: From soft sets to information systems. In: Hu, X., Liu, Q., Skowron, A., Lin, T.Y., Yager, R.R., Zhang, B. (eds.). Proceedings of Granular Computing, IEEE(2), pp. 617–621 (2005)
4. Sezgin, Aslıhan: Akın Osman Atagün: on operations of soft sets. Comput. Math. Appl. **61**, 1457–1467 (2011)
5. Maji, P.K. et al.: Fuzzy soft-sets. J. Fuzzy Math. **9**(3), 589–602 (2001)
6. Roy, A.R., Maji, P.K.: A fuzzy soft set theoretic approach to decision making problems.bib of. Comput. Appl. Math. **203**, 412–418 (2007)
7. Ahmad, B., Kharal, A.: On fuzzy soft sets. Adv. Fuzzy Syst. 1–6 (2009)
8. Xiao, Z., Gong, K., Zou, Y.: A combined forecasting approach based on fuzzy soft sets. J. Comput. Appl. Math. **228**, 326–333 (2009)
9. Shabir, M., Ali, M.I.: Comments on De Morgan's law in fuzzy soft sets. J. Fuzzy Math. **18**(3), 679–686 (2010)
10. Çağman, N., Çıtak, F., Enginoglu, S.: Fuzzy parameterized fuzzy soft set theory and its applications. Turk. J. Fuzzy Syst. **1**(1), 21–35 (2010)
11. Feng, F., Jun, Y.B., Liu, X., Li, L.: An adjustable approach to fuzzy soft set based decision making. J. Comput. Appl. Math. **234**, 10–20 (2010)
12. Tanay, B., Kandemir, M.B.: Topological structure of fuzzy soft sets. Comput. Math. Appl. **61**, 2952–2957 (2011)

13. Aktas, H., Cagman, N.: Soft sets and soft groups. Inf.Sci. **177**, 2726–2735 (2007)
14. Feng, F., Jun, Y.B., Zhao, X.: Soft semirings. Fuzzy Set. Syst.: Theor. Appl. 56(10), 2621–2628 (2008)
15. Jun, Y.B., Park, C.H.: Applications of soft sets in ideal theory of BCK/BCI-algebras. Inf. Sci. **178**, 2466–2475 (2008)
16. Jun, Y.B., Lee, K.J., Park, C.H.: Soft set theory applied to commutative ideals in Bck-algebras. J. Appl. Math. Inform. **26**(3–4), 707–720 (2008)
17. Park, C.H., Jun, Y.B., Öztürk, M.A.: Soft WS-algebras. Commun. Korean Math. Soc. **23**(3), 313–324 (2008)
18. Jun, Y.B., Kim, H.S., Neggers, J.: Pseudo d-algebras. Inf. Sci. **179**, 1751–1759 (2009)
19. Jun, Y.B., Park, C.H.: Applications of soft sets in Hilbert algebras. Iran. J. Fuzzy Syst. **6**(2), 55–86 (2009)
20. Jun, Y.B., Lee, K.J., Khan, A.: Soft ordered semigroups. Math. Logic Quart. **56**(1), 42–50 (2010)
21. Kazancı, O., Yılmaz, Ş., Yamak, S.: Soft sets and soft BCH-algebras. Hacet. J. Math. Stat. **39**(2), 205–217 (2010)
22. Ali, M.I., Shabir, M., Naz, M.: Algebraic structures of soft sets associated with new operations. Comput. Math. Appl. **61**, 2647–2654 (2011)
23. Zou, Y., Xiao, Z.: Data analysis approaches of soft sets under incomplete information. Knowl.-Based Syst. **21**, 941–945 (2008)
24. Maji, P.K., Roy, A.R.: An application of soft sets in a decision making problem. Comput. Math. Appl. **44**, 1077–1083 (2002)
25. Maji, P.K., Roy, R.: A fuzzy set theoretic approach to decision making problems. J. Comput. Appl. Math. Appl. **203**, 412–418 (2007)
26. Çağman, Naim, Enginoğlu, Serdar: Soft set theory and uni-int decision making. Eur. J. Oper. Res. **207**, 848–855 (2010)
27. Majumdar, P., Samanta, S.K.: On soft mappings. Comput. Math. Appl. **60**, 2666–2672 (2010)
28. Babitha, K.V., Sunil, J.J.: Soft set relations and functions. Comput. Math. Appl. **60**, 1840–1849 (2010)
29. Som, T.: On the theory of soft sets, soft relation and fuzzy soft relation. In: Proceedings of the National Conference on Uncertainty: A Mathematical Approach, UAMA-2006, pp. 1–9. Burdwan(India) 7–8 Sept (2006)
30. Luo, C.Z.: Introduction to fuzzy sets(1). Beijing Normal University Press, pp. 1–486 (1989) (in chinese)

Fire Detection in Video Using Fuzzy Pattern Recognition

Lin Wang, Aiguo Li, Xin Yao and Kaiqi Zou

Abstract An early fire flame video detection method based on multi-feature fusion of video flame is proposed in this paper. The background subtraction algorithm is used to extract the moving object. The suspected region is segmented by the statistical color model of flame. The jumping frequency feature of flame is extracted according to the motion direction code of object centroid. The ratio of square of perimeter to area of the suspected region is computed to extract circularity. The ratio of the total number of pixels in the suspected region to the area of the minimal enclosing rectangle of the suspected region is calculated to extract rectangularity. The number of sharp angles of flame is extracted according to the shape feature of sharp angles. The fire flame is judged by integrating four features by a fuzzy pattern recognition algorithm based on maximal membership degree principle. Experimental results show that this algorithm can detect early fire flame quickly and efficiently regardless of the scene changes of indoor and outdoor and several interference conditions by the test of shoot video and public video.

Keywords Fire flame video detection · Multi-feature fusion · Fuzzy pattern recognition · Jumping frequency · Circularity · Rectangularity · Sharp angle

1 Introduction

Fire is an unpredictable disaster, which is derived from uncontrolled combustion. It is necessary for us to notice early fire timely and give the alarm promptly when it occurs, in order to control early fire effectively. Classical fire detector such as photoelectric smoke detector and temperature detector can be easily affected by

L. Wang (✉) · A. Li · X. Yao
College of Information Science and Technology, Dalian Maritime University, Dalian 116026, China
e-mail: wanglin_dmu@163.com

K. Zou
College of Information Science, Dalian University, Dalian 116622, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_13

some environmental factor, such as area, dust, humidity, airflow. Thermal radiation and soot, produced by combustion, cannot diffuse to detection range of sensor in time when the distance between kindling point and sensor is relatively far. Hence, fire detection machine with conventional sensor is difficult to timely and effectively detect occurrence of fire in complex environment. In order to deal with early fire detection problem in large space, i.e. highway tunnel and stadiums, researchers have implemented fire detection by using image processing technique. Phillips et al. [1] practice a manual segmented color model and apply Gaussian smoothing to obtain color histogram. Based on the histogram and appropriate threshold value, Phillips obtain a Bool detection function of flame chromaticity. Huang [2] concludes statistical color distribution of flame from RGB and HSI color space, then recognizes flame with help of area increment in early fire combustion and kinetic stability. Celik [3] propose a flame color model based on RGB color space via investigating several flame image. With the research of pulse feature of flame, Yuan et al. [4] propose outline volatility model which is based on normalized Fourier descriptor and measure time-space jumping feature of flame. Zhang et al. [5] obtain difference features between flame objective and interference source by analyzing Fourier frequency spectrum of flame color. Ko et al. [6] propose a method of flame recognition using support vector machine classifier. Jenifer [7] uses probability statistics to obtain color, area, surface roughness, boundary roughness and gradient of flame, and then makes decision by Bayes classifier. Li et al. [8] proposes a fire monitoring mechanism with infrared camera, which can extract degree of irregularity, angular vector, jumping frequency and inter-frame correlation. They optimize training sample by combining subtractive clustering with fuzzy c-means and train neural network classifier of flame recognition in two different statuses.

In order to detect early fire flame timely and efficiently, an early fire detection method based on multi-feature fusion of video flame is proposed, which integrates the flame features of jumping frequency, circularity, rectangularity and the number of sharp angles, in this paper. Firstly, the motion region is extracted by background subtraction method and the candidate region is segmented by statistical color model of flame in RGB and HSI spaces. Then, the feature values of jumping frequency, circularity, rectangularity and the number of sharp angles in candidate region are extracted. Finally, the fuzzy pattern recognition algorithm, based on maximal membership principle, is used to detect flame for video under test. Experimental results show that this algorithm achieves better recognition rate than typical flame recognition algorithm under the condition of high light intensity and multiple disturbances in inside and outside.

2 Proposed Fire Detection Algorithm

A flowchart of the fire detection algorithm using video images is depicted in Fig. 1. The algorithm is composed of two phases: training and classification. In the training phase, the training data sets are extracted from the training videos and used to train

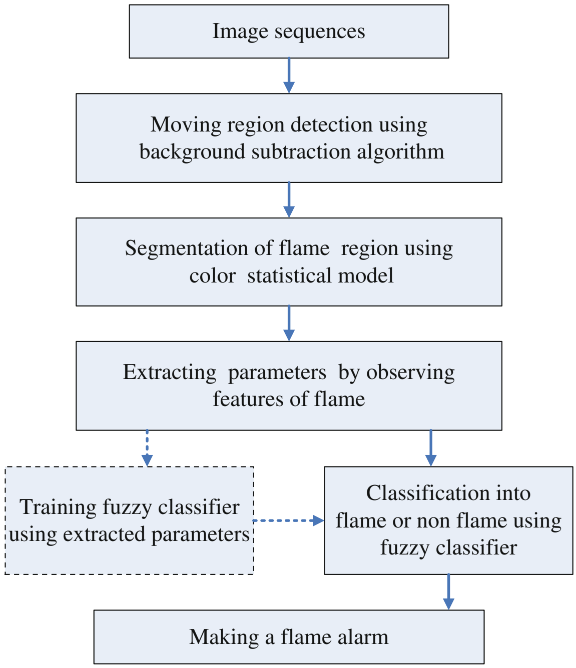


Fig. 1 Flowchart of the proposed flame-detection algorithm

the fuzzy classifier. In the classification phase, the trained fuzzy classifier is utilized to distinguish between flame and non-flame. In the following sections, the proposed flame detection algorithm is presented in detail.

2.1 Moving Region Detection Using Background Subtraction

Moving region detection is a fundamental task in video fire detection, and it is the first step of the proposed algorithm. A number of approaches have been proposed to detect moving regions in video images from static cameras: background subtraction, frames difference and optical flow. The background subtraction method is commonly used because of its simplicity and effectiveness. The method separates the foreground objects from the background in a sequence of video frames. Different methods in background subtraction have been proposed in which each method has different strengths and weaknesses in terms of performance and computation time. To detect a moving region, a modified hybrid background estimation method [9] is used. Suppose x_n is the intensity value at each spatial location k, l in the nth frame. The unit of intensity has values from 0 to 255. Assume that B_n is the background intensity value at the same pixel position in the previous frame, B_{n+1} is computed using the following where the weight value α is a positive real number between 0 and 1, and a symbol t is the disorder threshold to distinguish fire pixels, including moving pixels, from background pixels. The unit of threshold t is the pixel difference. If the frame difference is under the threshold t pixel, the same pixel in the next frame is linearly combined to the background pixel in the current frame. Meanwhile, if the frame difference is over the threshold t pixel, the pixel at location k, l is a declared candidate fire pixel and it used to replace the same pixel in the next background frame. In the paper, we set t as 4 pixels which produced the best results during experimentation.

$$B_{n+1} = \begin{cases} if\,|B_n[k,l] - x_n[k,l]| < t \\ then\ \ \alpha B_n[k,l] + (1-\alpha)x_n[k,l] \\ else \\ x_n[k,l] \end{cases} \tag{1}$$

2.2 Segmentation of Suspected Region Using Fire Color Model

In this paper, the statistical color model of flame [10] (2), (3) are used to segment the suspected region of flame.

$$\begin{cases} R > R_T \\ R \geq G > B \\ S \geq (255 - R)\ S_T/R_T \end{cases} \tag{2}$$

$$0.20 \leq \frac{B}{G+1} \leq 0.60 \tag{3}$$

where, R_T is threshold of red component in RGB color space, and S_T is threshold of hue saturation in HSI color space, the value range from 115 to 135 for R_T, the value range over 200 for S_T. Hue saturation value of flame is greater than Hue saturation value of its reflected light. The unsaturated reflected light around suspected flame objective is excluded by calculating saturation value. The foreground images of moving flame and moving vehicle lighting extracted by statistical color model algorithm and corrected by mathematical morphology are shown in Fig. 2.

Fig. 2 Foreground images of moving flame and moving vehicle lighting using background subtraction and statistical color model. **a** Flame image. **b** Foreground image of flame. **c** Vehicle lighting image. **d** Foreground image of moving vehicle lighting

2.3 Parameters Extraction

The selected candidate regions of fire, shown in Fig. 2b, d, can be fire or non-fire. By investigating several video images, we observed that the temporal and spatial features of flame vary during the manifestation of fire. The nature of the varying features of fire flame is utilized in this study to detect fire. Thus, we extract the following features of the candidate regions to identify either fire or non-fire.

2.3.1 Jumping Frequency

Flame is jumping when fire occurs. The jumping frequency of flame is stable in different environment. The flame jumping frequency in time domain can be extracted as the important feature of early fire flame [11]. The moving direction of the object is divided into 8 sections, 1, 2, 3, 4, 5, 6, 7, 8 represent direction code for each section respectively, which indicates angular range, i.e. 0°–45°, 45°–90°, 90°–135°, 135°–180°, 180°–270°, 270°–315°, 315°–360°.

The centroid coordinate of the candidate object in (*i*-1)-th and *i*-th frame of image are denoted as $(x_{i1}y_{i1})$ and $(x_i y_i)$. The motion direction code θ_i in *i*-th frame is estimated according to the following rules.

When $x_m > y_m \geq 0$, $\theta_i = 1$; when $y_m \geq x_m > 0$, $\theta_i = 2$; when $y_m > -x_m \geq 0$, $\theta_i = 3$;

When $-x_m \geq y_m > 0$, $\theta_i = 4$; when $x_m < y_m \leq 0$, $\theta_i = 5$; when $y_m \leq x_m < 0$, $\theta_i = 6$;

When $-y_m > x_m \geq 0$, $\theta_i = 7$; when $x_m \geq -y_m > 0$, $\theta_i = 8$; when $y_m = x_m = 0$, $\theta_i = 0$.

Where, $x_m = x_i - x_{i-1}$, $y_m = y_i - y_{i-1}$.

To obtain the major motion direction of object, the motion direction code of object is accumulated in the successive frames.

$$A_\theta = \sum_{i=0}^{n-1} H(\theta) \tag{4}$$

where, θ = 1, 2,…, 8. A_θ represents occurrence number of direction code θ, $H(\theta)$ indicates accumulation consequence of θ if the direction code of objective is θ in *i*-th frame of image.

In this paper, algorithm recognizes jumping feature of centroid in flame objective region based on the model of jumping frequency in time domain, as formula (5) shows.

$$f = \frac{F_{FPS}(A_\theta + A_{\theta+1} + A_{\theta+4} + A_{\theta+5})}{2N} \tag{5}$$

where, f is jumping frequency of centroid, N is sum of tested frame, F_{FPS} is frame rate of video, θ123, in this paper, θ2, A_θ is occurrence number of direction code θ in motion direction section.

2.3.2 Circularity Feature

The edge outline of early fire flame is irregular while shape of interfering illuminant is regular, such as lamplight of moving vehicle lighting and flashlight. Therefore, the circularity can be used as a feature value of early fire flame to distinguish early flame from other suspected interfering source.

The formula of circularity is defined as

$$C = \frac{4\pi A}{P^2} \tag{6}$$

where, P and A are perimeter and area of foreground objective. If object region is circle, $C1$, otherwise, $0 < C < 1$. Edge outline of uncontrolled flame is extreme irregular, whose circularity is comparatively small.

2.3.3 Rectangle Filling Coefficient

The rectangle filling coefficient is used to measure the edge irregularity of early fire flame. Therefore, the rectangle filling coefficient is a feature of early fire flame recognition.

The rectangle filling coefficient of the object is defined as

$$R = \frac{A_0}{A_R} \tag{7}$$

where, A_0 is the area of the object, A_R is the area of its minimum enclosing rectangle. If the object is a rectangle, $R1$, otherwise, $0 < R < 1$.

2.3.4 The Sharp Angle Feature of Flame

The jitter edge of early fire flame leads to more sharp angles of flame in single frame image [12]. However, the number of sharp angles of moving car light, flashlight light, controlled candle flame is relatively small. Therefore, the number of sharp angles can be used as a feature to recognize early fire flame.

The shape of flame sharp angle is represented as irregular slender triangle in image. Firstly, to find suspected sharp angle vertex in the current frame, in this paper. Then, to construct the suspected acute triangle by suspected sharp point and 40 points around the left and right neighborhood of the suspected sharp vertex, and to select the ordinate of 5-th, 15-th and 40-th point on the left and right of suspected sharp vertex for to calculate its width $W_1 W_2 W_3$, respectively. If $W_1 < W_2 < W_3$, the angle is sharp angle. Finally, the value of three edge length a, b, c on the left, right and low is calculated based on vertex coordinate of three angle of suspected acute angle triangle. According to Heron's formula $sp(pa)(pb)(pc)$, $p(abc)2$ and $scH2$, the height H on the bottom edge c of triangle is calculated. The acute angle is accord with feature of flame acute angle if the height is greater than certain threshold. Then the number of acute angles is obtained in the current frame.

2.4 Flame Alarm Decision Using a Fuzzy Classifier

To classify a candidate pixel as fire or non-fire pixel, a fuzzy classifier is used. Let U represent the given entire collection of objects to be identified, every object u of U has p characteristic indexes u_1, u_2,…, u_p. Every characteristic index represents a certain feature of u. So p characteristic indexes can determine an object $u = (u_1, u_2, \ldots, u_p)$, which is called characteristic vector. U can be divided into n categories and every category is one fuzzy set of U. The fuzzy set can be denoted as A_1, A_2,…, A_n, which is called fuzzy pattern.

Fuzzy pattern recognition method is attributing the object u to its similar category A_i.

When a recognition algorithm is applied to an object u, a set of membership $A_1(u)A_2(u)\ldots A_n(u)$ are generated which represents the degree of u belonging to A_1, A_2,…, A_n, respectively. After the establishment of the membership function group of fuzzy pattern, the object u can be determined to which category it belongs in accordance with a certain membership principle.

For fire recognition, the characteristic vector is composed of the circularity, rectangle filling coefficient, the number of sharp angle and jumping frequency. The circularity of suspected region is computed in the current frame, if the circularity is greater than the threshold T_{C1} and less than T_{C2}, which means that the region satisfies the circularity feature, this frame is accumulated. Otherwise, the frame is excluded. The rectangle filling coefficient of suspected region is computed in the current frame. If the rectangularity is greater than the threshold T_{R1} and less than T_{R2}, this frame is accumulated. Otherwise, the frame is excluded. The flame sharp angle of suspected region in the current frame is recognized according to the shape feature of flame sharp angle image. The number of sharp angles in each frame can be obtained. If the number of sharp angles is not less than T_{SAN}, this frame accumulates. Otherwise, it is excluded. The motion direction code of centroid in suspected region can be collected in each testing period. The centroid jumping frequency of the main motion direction can be computed according to expression (5). If the jumping frequency is not less than T_{JF}, this region satisfies flame jumping frequency feature. Otherwise it cannot satisfy. The fuzzy pattern recognition process is shown in Fig. 3.

In sample video, the features of clear flame object are detected according to threshold of circularity, rectangularity and the number of sharp angles. The ratio of accumulated value of each feature to the total number of frames in a detection period is calculated. The mean template can be obtained. The mean template of interference object which commonly appears in video is calculated and the template

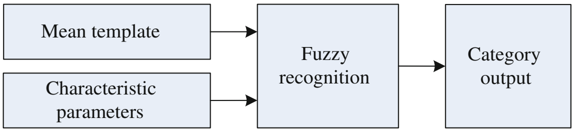


Fig. 3 Fuzzy recognition process

Table 1 Feature mean template values of flame and interference object

Pattern category	a_1	a_2	a_3	a_4
Flame object A_1	9.40	0.84	0.92	0.08
Flame object A_2	8.50	0.82	0.86	0.50
Interference object A_3	3.20	0.74	0.65	0.08
Interference object A_4	2.20	0.50	0.50	0.04

value can be obtained. The template feature parameters of flame and interference object are given in Table 1.

The objects are divided into several categories for recognition according to maximal membership principle in this paper. The feature membership of candidate flame object is calculated and pattern classification recognition is carried out. If the recognition result is flame, an alarm signal is given. Otherwise, frame counter is cleared.

3 Experimental Results and Analysis

Some sample videos in this paper are self-shoot video and the others are public video on the Internet. The algorithm is implemented on Visual C++ 6.0. The parameter definition of experiment is given in Table 2.

The jumping frequency, circularity, rectangularity and the number of sharp angle of flame sample are extracted in this paper. Their experimental data are shown in Fig. 4.

In this paper, the flame and non-flame test videos in the conditions of different environments and disturbances, as shown in Figs. 5 and 6, are experimentally tested, and then classified according to fuzzy pattern recognition based on maximal membership principle. The experimental data are presented in Tables 3 and 4. In Fig. 5, video1, 2 and 3 are fire flame simulation video of paper, gasoline and wood combustion respectively under indoor and outdoor with light and different environments. The recognition rate of the algorithm in this paper for video 1, 2 and 3 are more than 92 %. Video 4 is the flame video of paper combustion shooting outdoors at night. The flame recognition rate for video 4 is 90 % because of empty-illumination. In Fig. 6, video 1 and 2 are the driving videos in highway

Table 2 Parameter definition

T_{C1}	T_{C2}	T_{R1}	T_{R2}	T_{SAN}	T_{JF}	N
0.12	0.45	0.35	0.76	3	6.8	48

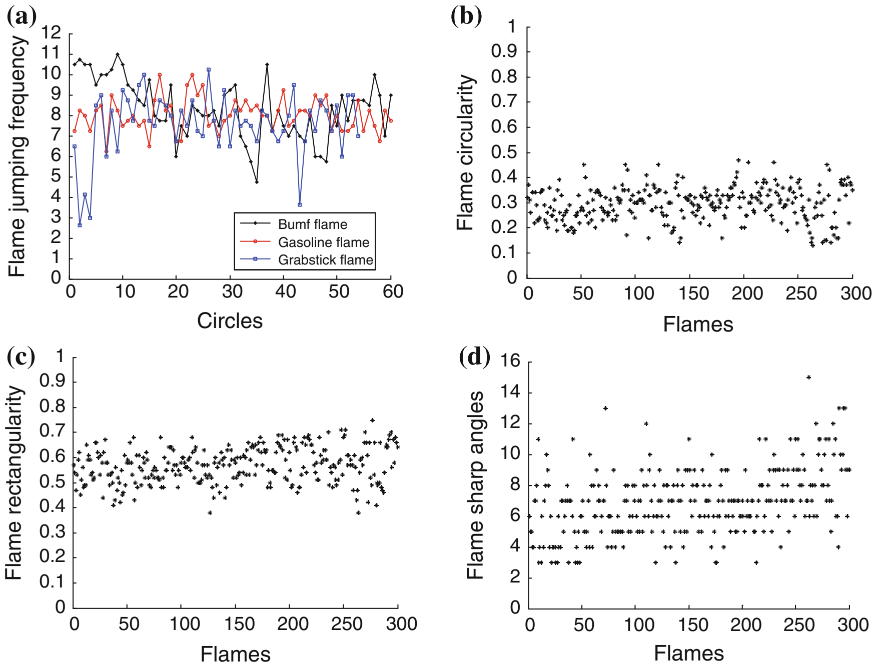


Fig. 4 Distribution of flame sample feature data. **a** Distribution of flame jumping frequency. **b** Distribution of flame circularity. **c** Distribution of flame rectangularity. **d** Distribution of flame sharp angles

Fig. 5 Video images of flame samples. **a** Video 1, **b** video 2, **c** video 3, **d** video 4

Fig. 6 Video images of non-flame samples. **a** Video 1, **b** video 2, **c** video 3, **d** video 4

Table 3 Experimental data of flame videos

Video name	Light radiation	Detection period	Alarm times	Missing alarm times
Video 1	Strong	18	18	0
Video 2	Strong	4	4	0
Video 3	Weak	54	50	4
Video 4	No	11	10	1

Table 4 Experimental data of interference videos

Video name	Light radiation	Detection period	Alarm times	False alarm times
Video 1	Strong	8	0	0
Video 2	Strong	54	3	3
Video 3	Weak	56	4	4
Video 4	Weak	38	5	5

tunnel in the conditions of strong light. The recognition rate of the algorithm for car light can be more than 94 %. Video 3 and 4 are driving videos shooting roadside at night for flame interference simulation. Due to the weaker light of driving road in video 3 and 4, the recognition rates of the proposed algorithm are 93 % and 87 % respectively. Experimental results show that, the proposed algorithm can achieve early fire flame recognition in the conditions of indoor and outdoor with lights and different disturbances.

4 Conclusion

In this paper, an early fire flame video detection method based on multi-feature fusion is proposed. Firstly, in every detection period, the background subtraction algorithm is used to extract moving object and the statistical color model of flame is used to segment suspected region. Then, three feature values of frame in image sequences are extracted by circularity, rectangularity and the number of sharp angles of flame. The jumping frequency of flame centroid is calculated according to the motion direction codes of object centroid. Finally, the early fire flame is recognized by a fuzzy pattern recognition algorithm based on maximal membership degree principle. Experimental results show that the algorithm can recognize early fire flame effectively under the condition of high light intensity and multiple disturbances in indoor and outside. However, it is possible that the algorithm neglect or misrecognize flame image under the special condition of empty illumination or weak light intensity. The algorithm still needs to be further improved and optimized for stronger robustness.

Acknowledgements This paper was supported by the Fundamental Research Funds for the Central Universities (Grant No. 313 2014 321, 313 2014 301) and the Project of Applied Basic Research Project of Ministry of Transport (Grant No. 2015 319 225 210).

References

1. Phillips III, W., Sha, M., da Vitoria Lobo, N.: Flame recognition in video. In: Proceedings of the Fifth IEEE Workshop on Applications of Computer Vision, IEEE, pp. 224–229 (2000)
2. Huang, P.H., Su, J.Y., Lu, Z.M., Pan, J.S.: A Fire-alarming method based on video processing. In: International Conference on Intelligent Information Hiding and Multimedia, Institute of Electrical and Electronic Engineering Computer Society, pp. 359–364. Piscataway, NJ, USA (2006)
3. Celik, T., Demirel, H., Ozkaramanli, H., Uyguroglu, M.: Fire detection using statistical color model in video sequences. J. Vis. Commun. Image Represent. **18**(2), 175–176 (2007)
4. Yuan, F.N., Liao, G.X., Zhang, Y.M., Liu, Y., Yu, C.Y., et al.: Feature extraction for computer vision based fire detection. J. Univ. Sci. Technol. China **36**(01), 39–43 (2006)
5. Zhang, J.H., Zhuang, J., Du, H.F., et al.: A Flame Detection Algorithm Based on Video Multi Feature Fusion, pp. 784–792. Springer Verlag, Xian, China (2006)
6. Ko, B.C., Cheng, K.H., Nam, J.Y.: Fire detection based on vision sensor and support vector machines. Fire Saf. J. **44**(3), 322–329 (2009)
7. Jenifer, P.: Effective visual fire detection in video sequences using probabilistic approach. emerging trends in electrical and computer technology (ICETECT). In: 2011 International Conference, pp. 870–875. IEEE Computer Society, Piscataway, NJ, USA (2011)
8. Li, W.H., Wang, Y., Zhao, R.L.: Fire detection with day-night infrared camera. J. Hunan Univ. (Nat. Sci.) **39**(8), 73–80 (2012)
9. Collins, R.T., Lipton, A.J., Kanade, T.: A system for video surveillance and monitoring. In: Proceedings of the American Nuclear Society(ANS) Eighth International Topical Meeting on Robotics and Remote Systems, pp. 1–15. Pittsburgh, PA (1999)
10. Chen, T.H., Wu, P.H., Chiou, Y.C.: An early fire-detection method based on image processing. In: Proceedings of IEEE International Conference on Image Processing IEEE, pp. 1707–1710. Singapore (2004)
11. Wang, L., Li, A.G., Hao, C.Y.: A fire detection method using the flame image jumping feature. In: Proceedings of the 33rd Chinese Control Conference, pp. 7421–7425. Shanghai Systems Science Publisher, Nanjing, China, (2014)
12. Wu, L.B., Song, W.G., Lu, J.C.: A new fire recognition method in image fire monitoring. Fire Saf. Sci. **6**(2), 60–66 (1997)

Fuzzy Prediction in Classification of AdaBoost Algorithm

Hong-yang Zhu, Yi Ding, Hong Gao and Wei Liu

Abstract To resolve the question that traditional classification algorithm can't provide a result with abundant information, using the method of theoretical analysis and experimental surveys, a new AdaBoost algorithm which expresses the results as membership degree is proposed. Combining the Fuzzy mathematical model and membership function, it finds the crucial feature to the real label as the basis of weak classification algorithm and classifies things according to the membership function. the prediction of classical classification algorithm about certain category, "yes" or "no", is replaced by the membership degree of Fuzzy AdaBoost with a prediction which is kind of more reference and value as well as the higher accuracy.

Keywords Adaboost · Fuzzy mathematics · Classification algorithm · Membership function · Fuzzy prediction · Ensemble learning

1 Introduction to AdaBoost

Nowadays, it is information society that makes differences both in our daily life and professional fields. There are more sources and export of information data than ever before. It is an important process to classify these information data into the correct category to take full advantage of them. It can get the most values from some object as only that it has been put in the right position. Some excellent classification algorithms have been risen in response to the proper time and conditions. Some among them can get a categorized result with an equal, if not higher, accuracy as manual work.

For binary classification, the categorized results of traditional classification algorithms are expressed as "Yes" or "No" about a certain category. In other words,

H. Zhu · H. Gao · W. Liu
Department of Mathematics, Dalian Maritime University, Dalian 116026, China

Y. Ding (✉)
Engineering College of Guangdong Ocean University, Zhanjiang 524088, China
e-mail: sky.pai@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_14

traditional algorithms define object with a label which is either "+1" or "−1". It can't furnish valuable enough reference sometimes, so a new AdaBoost algorithm, the Fuzzy AdaBoost algorithm, combining the Fuzzy mathematical model and membership function is proposed which expresses the results as membership degree.

There is a general conflict in classification between the high accuracy and the difficulty in designing an accurate strong learner. AdaBoost Based on the ensemble learning, has resolved this conflict. It designs a new weak learner with current weight and assigns weights for the training samples again in each training iteration. After all, it combines those weak learners into a strong one with a much higher accuracy [1, 2].

Training set: $S = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_i, y_i)\}, x_i \in X, y_i \in \{+1, -1\}$, $i = 1, 2, 3, \ldots, m$

Processes of traditional AdaBoost is as follow [3, 4]:

(1) initialize the weight of training set: $\omega_1(i) = \frac{1}{m}$;
(2) iteration: $t = 1, 2, 3, \ldots, T$:
(2.1) construct a weak classifier h_t with current weight distribution ω_t;
(2.2) compute the error rate ε_t of the weak classifier h_t, and the weight of h_t:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right); \tag{1}$$

(2.3) reassign weights for the training samples [5, 6]:

$$\omega_{t+1}(i) = \frac{\omega_t(i) \cdot e^{(-\alpha_t \cdot y_i \cdot h_t(x_i))}}{Z_t} \tag{2}$$

$Z_t = \sum_{i=1}^{m} \omega_t(i) \cdot e^{(-\alpha_t \cdot y_i \cdot h_t(x_i))}$, is the normalization. Weights of error-classifying samples will increase if the error rate is below 50 %.

(3) the strong leaner:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \tag{3}$$

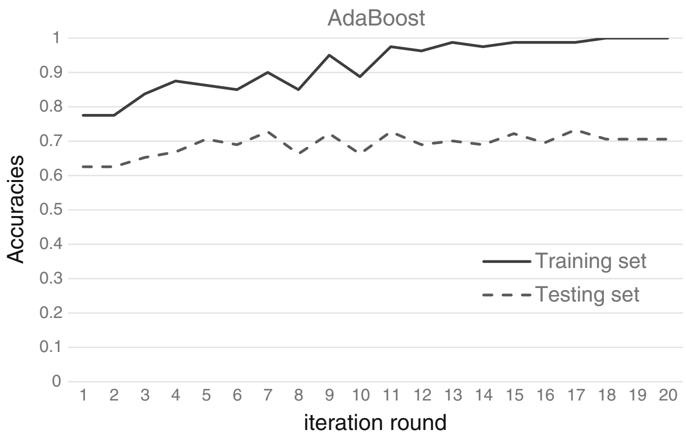


Fig. 1 Accuracy of AdaBoost in each iteration

2 Fuzzy Mathematics

2.1 Sets

Research object is usually restricted in appropriate range which is called domain or universe. Every object in the domain is called element; the entirety of Elements with the same attributes is called subset of domain. For the traditional set theory, in the domain U, the relation between a random element x and a random subset A is requested as only two possibility: x belongs to A, denoted as $x \in A$; x does not belong to A, denoted as $x \notin A$ [7].

If A is a subset of domain $U(A \subseteq U)$, Eigen function of A, $\chi_A(x)$, is the mapping from U to $\{0,1\}$, denoted as $\chi_A(x) : U \to \{0,1\}, x \mapsto \chi_A(x)$, and

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}, \forall x \in U \tag{4}$$

2.2 Fuzzy Subsets and Membership Function

L.A.Zadeh extended the value area of Eigen function from $\{0,1\}$ to the interval of $[0,1]$ as well as promoted the concept of membership function. Fussy subsets have been defined as follow: A is the mapping from U to $[0,1]$, namely [8],

$$A : U \to \{0,1\}, x \mapsto A(x). \tag{5}$$

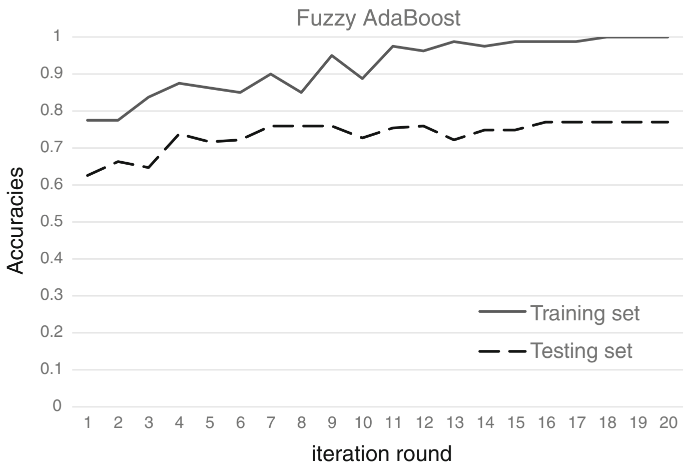


Fig. 2 Accuracy of fuzzy AdaBoost in each iteration

A is called the fussy subset of domain U; Function $A(\cdot)$ is called the membership function of A; $A(x)$ is called the fuzzy membership of x regarding A [9, 10].

2.3 Fuzzy Mathematical Model

Membership function of fuzzy subsets structured in real number domain can describe the degree that object belongs to certain category objectively in fussy mathematic [11].

Minor type and major type of K parabolic distribution and Cauchy distribution are used to ascertain membership function in this paper.

K parabolic distribution:

The minor type

$$A(x) = \begin{cases} 1, & x < a \\ (\frac{b-x}{b-a})^2, & a \le x \le b\,; \\ 0, & b < x \end{cases} \tag{6}$$

The major type

$$A(x) = \begin{cases} 0, & x < a \\ (\frac{x-a}{b-a})^2, & a \le x \le b\,. \\ 1, & b < x \end{cases} \tag{7}$$

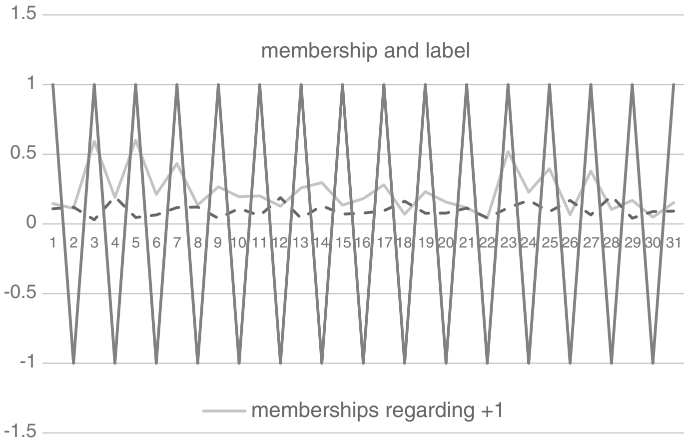


Fig. 3 The fuzzy prediction of fuzzy AdaBoost

Cauchy distribution:
The minor type

$$A(x) = \begin{cases} 1, & x \leq a \\ \frac{1}{1+(x-a)^2}, & a < x \end{cases}; \tag{8}$$

The major type [12]

$$A(x) = \begin{cases} \frac{1}{1+(x-a)^2}, & x \leq a \\ 1, & a < x. \end{cases} \tag{9}$$

3 Fuzzy AdaBoost

Traditional AdaBoost can get an accurate strong classifier which can satisfies the necessary of application based on a weak learning algorithm whose central theme is easy to understand and operate. But the prediction result of traditional AdaBoost can only provide a label, "+1(Yes)" or "−1(No)", whose information is not integrated and systematic.

The fuzzy predictions composed by memberships regarding both categories are more valuable and accurate by integrating traditional AdaBoost with fussy mathematics theory.

3.1 Training

The weak classifier of fuzzy AdaBoost is the best threshold of the crucial feature to the real label under current weight distribution.

Assumed that samples have N features, the process is as follow:

(1) initialize the weight of training set: ω_1;

(2) iteration: $t = 1, 2, 3, \ldots, T$;

(2.1) sort samples according to every feature under current weight distribution ω_t, then find the most well-organized one n_{t}. For example, there are some cancer cells and normal cells. It finds that the bigger ones are usually cancer cells while the smaller ones are usually normal cells after sorting the cells according to their volumes, concentration of cytosol and any other character, so the volume is the crucial feature to the real label;

(2.2) find the optimal threshold h_t of n_t, and classify objects whose values of feature n_t are greater than h_t into the same category and the ones lesser than h_t into the other category that makes the error ε_t minimal. Denote $\alpha_{\mathrm{t}} = \frac{1}{2}\ln(\frac{1-\varepsilon_t}{\varepsilon_t})$ as the weight of the weak classifier h_t. The weight of h_t is coincident with the accuracy, so if the accuracy of h_t is less than50 %, the weight is minus;

(2.3) it may make mistakes while h_t is regarded as a threshold, so find the values $a_t \le h_t$ and $b_t \ge h_t$ resulting in that the all the samples whose feature values of n_t are lesser than a_t belong to "+1" (or "−1") and the ones greater than b_t to the other category. Denote $< a_t, h_t >$ as the fuzzy interval of "+1"(or"−1"), $< h_t, b_t >$ as the fuzzy interval of the other.

(3) reassign weights for the training samples according to the correctness of h_t: $\omega_{t+1}(i) = \frac{\omega_t(i) \cdot e^{(-\alpha_t \cdot y_i \cdot h_t(x_i))}}{Z_t}$, $Z_t = \sum_{i=1}^{m} \omega_t(i) \cdot e^{(-\alpha_t \cdot y_i \cdot h_t(x_i))}$, is the normalization.

3.2 Predictions

The membership of objects regarding "+1", $A_t(+1)$, and "−1", $A_t(-1)$ can be ciphered out as there are the fuzzy interval of "+1" and "−1" in each iteration according to the membership function of K parabolic distribution and Cauchy distribution (while $a_t = h_t$ or $b_t = h_t$):

The minor type

$$A_t(x) = \begin{cases} 1, & x < a_t \\ (\frac{h_t - x}{h_t - a_t})^2, & a_t \le x \le h_t \\ 0, & h_t < x \end{cases} \tag{10}$$

or

$$A_t(x) = \begin{cases} 1, & x \le a_t \\ \frac{1}{1+(x-a_t)^2}, & a_t < x(a_t = h_t) \end{cases}; \quad (11)$$

the major type

$$A_t(x) = \begin{cases} 0, & x < h_t \\ (\frac{x-h_t}{b_t-h_t})^2, & h_t \le x \le b_t \\ 1, & b_t < x \end{cases} \quad (12)$$

or

$$A_t(x) = \begin{cases} \frac{1}{1+(x-b_t)^2}, & x \le b_t(b_t = h_t) \\ 1, & b_t < x \end{cases}. \quad (13)$$

Normalize the weights of each membership: $\lambda_t = \frac{\alpha_t}{\sum_{i=1}^{T} \alpha_i}$, so the final membership of object regarding "+1", $K(+1) = \sum_{i=1}^{T} \lambda_i A_i(+1)$, and "−1", $K(-1) = \sum_{i=1}^{T} \lambda_i A_i(-1)$ can be ciphered out.

Fuzzy predictions can not only provide the memberships of object but also the classification by comparing the final memberships regarding "+1", $K(+1)$ and "−1", $K(-1)$. If $K(+1) > K(-1)$, this object will be classify into "+1", or else into "−1".

4 Data Experiment

Program of fuzzy AdaBoost is realized with MATLAB R2014a and a deal of experimental data indicates that fuzzy AdaBoost is ascendant with the small samples sets. So choose the spectf heart data set from UCI to experiment and analyze which contains 267 samples and 44 features and regard 80 samples among as training set, the others testing set.

Figures 1 and 2 illustrate the changes of training error and testing error of traditional AdaBoost and fuzzy AdaBoost on spectf heart data set as the number of iterations increases. Certainly, training accuracies of traditional AdaBoost are identical with fussy AdaBoost by the same training method, and obviously, testing accuracies of fuzzy AdaBoost are much higher and more convergent than that of traditional AdaBoost.

Figure 3 Corresponds the membership regarding "+1" and "−1" with the real label of 31 objects selected from the testing set of spectf heart data set.Black solid line in Fig. 3 means the real labels of these 31 objects, gray solid line means the memberships regarding "+1", and black dotted line means the memberships regarding "−1". If a value on the gray solid line is greater than the corresponding one on black dotted line, this object will be classified into "+1" or else "−1". It is revealed in Fig. 3 that the growth trend of gray solid line is consistent with Black solid line and that of black dotted line is opposite.

5 Conclusion

AdaBoost algorithm avoids the difficulty in finding the accurate strong learning algorithm tactfully with classifier combination. And fuzzy AdaBoost taking the crucial feature to the real label as the basis of weak classification algorithm replaces traditional simple prediction by fuzzy prediction composed by memberships. It provides more values and functionalities, especially for the small sample set, the result deduced by fuzzy prediction of fuzzy AdaBoost is more accurate and convergent in the testing set.

References

1. Xu, J., Wu, Q., Zhang, J., et al.: Exploiting universum data in AdaBoost using gradient descent [J]. Image Vis. Comput. **32**(8), 550–557 (2014)
2. Schapire, R.E., Singer, Y.: Improved boosting algorithms using confidence-rated predictions [J]. Mach. Learn. **37**(3), 297–336 (1999)
3. Wu, B., Ai, H., Huang, C., et al.: Fast rotation invariant multi-view face detection based on real adaboost[C]. In: Automatic Face and Gesture Recognition. Proceedings. Sixth IEEE International Conference on. IEEE, 2004, pp. 79–84 (2004)
4. Ying, C., Qi-Guang, M., Jia-Chen, L.I.U., et al.: Advance and prospects of AdaBoost algorithm[J]. Acta Automatica Sinica **39**(6), 745–758 (2013)
5. Sun, J., Jia, M.Y., Li, H.: AdaBoost ensemble for financial distress prediction: an empirical comparison with data from Chinese listed companies[J]. Expert Syst. Appl. **38**(8), 9305–9312 (2011)
6. Mathanker, S.K., Weckler, P.R., Bowser, T.J., et al.: AdaBoost classifiers for pecan defect classification[J]. Comput. Electron. Agric. **77**(1), 60–68 (2011)
7. Běhounek, L., Cintula, P.: From fuzzy logic to fuzzy mathematics: a methodological manifesto [J]. Fuzzy Sets Syst. **157**(5), 642–646 (2006)
8. Trillas, E., Termini, S., Moraga, C.: A naïve way of looking at fuzzy sets[J]. Fuzzy Sets Syst. 2014
9. Cao, H., Jia, L., Si, G., et al.: A clustering-analysis-based membership functions formation method for fuzzy controller of ball mill pulverizing system[J]. J. Process Control **23**(1), 34–43 (2013)
10. Marchler-Bauer, A., Lu, S., Anderson, J.B., et al.: CDD: a conserved domain database for the functional annotation of proteins.[J]. Nucleic Acids Res. **39**(2), D225–9 (2011)
11. Qian, G., Wang, H., Feng, X.: Generalized hesitant fuzzy sets and their application in decision support system[J]. Knowl.-Based Syst. **37**, 357–365 (2013)
12. Yao, K., Ke, H.: Entropy operator for membership function of uncertain set[J]. Appl. Math. Comput. **242**, 898–906 (2014)

Einstein Choquet Integral Operators for PROMETHEE II Group Decision Making Method with Triangular Intuitionistic Fuzzy Numbers

Lanlan Li, Lidong Wang and Binquan Liao

Abstract In order to reflect the interactions phenomena among attributes involved in decision making and simulate the mechanism of human's fuzziness thinking, a group decision model is established, in which the evaluation values are represented by triangular intuitionistic fuzzy numbers. By virtue of fuzzy Einstein Choquet integral geometric operator, the values of each expert with respect to multi-attributes are aggregated into triangular intuitionistic fuzzy numbers, and some interesting properties of that are also studied. PROMETHEE II model is employed to obtain group opinion, and select the best alternative. The proposed method can take full advantage of interactions information contained in the original data and the subjective information of experts. Moreover, an illustrative example is employed to demonstrate the practicality and effectiveness of the proposed approach.

Keywords Einstein operator · PROMETHEE II · Triangular intuitionistic fuzzy numbers · Choquet integral

1 Introduction

Multi-Attribute Decision Making (MADM) is to find the best from all the feasible alternatives with respect to a limited set of attributes, which is a common activity and plays an important role in a variety of practical fields. The decision process of MADM involved multiple decision makers is the multiple attribute group decision making (MAGDM), which can reduce the influences of subjective factors driven from the individual knowledge structure, professional skill and other aspects. Now, MAGDM has been applied to various fields, such as supplier selection [10], emergency alternative assessment [8] and evaluating the flexibility in a manufacturing system [2] and so on.

L. Li · L. Wang (✉) · B. Liao
Department of Mathematics, Dalian Maritime University, Dalian 116026, China
e-mail: ldwang@hotmail.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_15

In traditional MADM problems, decision makers assess each alternative under each attribute expressing with numerical values. As the inherent complexity and uncertainty in the decision making process, numerical values are not enough to reflect the human cognitive mechanism. Various fuzzy sets have been employed in MADM, including fuzzy sets, intuitionistic fuzzy sets, interval-valued intuitionistic fuzzy sets and uncertain linguistic variables. Triangular intuitionistic fuzzy number (TIFN) introduced by Shu et al. [11], which characterizes the uncertainty by a triangular fuzzy number to express decision information in different dimensions and avoid losing decision preference information. For this reason, TIFN has become an effective tool to cope with uncertainty and vagueness originating from imprecise knowledge or information in real applications. Now, TIFN has been generalized to other forms, such as intuitionistic trapezoidal fuzzy numbers [14], intuitionistic trapezoidal fuzzy linguistic numbers [5], interval-valued intuitionistic trapezoidal fuzzy numbers [18] and so on.

In MAGDM, aggregation operator is a crucial step to integrate evaluation values with respect to different attributes. Wang and Zhang [14] defined intuitionistic trapezoidal fuzzy numbers and their operational laws and proposed intuitionistic trapezoidal fuzzy weighted arithmetic averaging operator and weighted geometric averaging operator. Wan [15] defined new operation laws for trapezoidal intuitionistic fuzzy numbers (TrIFNs) and extended average operator of real numbers to four kinds of power average operators of TrIFNs, involving the power average operator of TrIFNs, the weighted power average operator of TrIFNs, the power ordered weighted average operator of TrIFNs, and the power hybrid average operator of TrIFNs. Wu [18] studied interval-valued intuitionistic trapezoidal fuzzy weighted geometric (IVITFWG) operator, the interval-valued intuitionistic trapezoidal fuzzy ordered weighted geometric (IVITFOWG) operator, and the interval-valued intuitionistic trapezoidal fuzzy hybrid geometric (IVITFHG) operator for MAGDM. In the real problems, there exist interactions phenomena between criteria [3, 4]. To some extent, that may be complementary or overlapping instead of independent of each other. Therefore, the aggregation operators based on non-additive measures is a more suitable tool than linear operators based on additive measures. Fuzzy Choquet integral operator is characterized by non-monotonic fuzzy measures, which is an effective tool to reflect the interactions. Wei et al. developed two intuitionistic fuzzy Choquet integral aggregation operators to reflect the correlation among the attributes or their ordered positions [17]. Beg and Rashid established a decision method by developing Choquet integral-based distance between trapezoidal-valued intuitionistic fuzzy values in classical TOPSIS method [1]. Patrick Meyera and Marc Roubensb performed the process of aggregation through the use of a fuzzy extension of the Choquet integral [9].

In this paper, the main object is to establish group PROMETHEE II method with intuitionistic fuzzy Einstein Choquet integral operators. Inspired by the references [19, 20, 22], the aim of this paper is to enrich triangular intuitionistic fuzzy numbers by investigating information aggregation methods utilizing Einstein Choquet integral operators when the decision information takes the form of triangular intuitionistic fuzzy numbers. Compared with the similar methods, the proposed method in this paper can maximize the group utility and minimize the individual regret simultaneously, which make the decision result more reasonable. The rest of the paper is organized as follows. In Sect. 2, some concepts of PROMETHEE II method and triangular intuitionistic fuzzy numbers are recalled. Section 3 presents triangular intuitionistic fuzzy Einstein Choquet integral geometric operator and its properties. Section 4 describes the proposed approach. In Sect. 5, one example of emergency classification is employed to demonstrate the practicality and effectiveness of the proposed approach. Finally, conclusions are drawn.

2 Preliminaries

In this section, some related concepts of triangular intuitionistic fuzzy numbers are recalled.

2.1 *Triangular Intuitionistic Fuzzy Numbers*

Definition 1 [21] Let $\tilde{a} = \langle (\underline{a}, a, \overline{a}) ; u_{\tilde{a}}, v_{\tilde{a}} \rangle$ be a triangular intuitionistic fuzzy number (TIFN), its membership function is defined as

$$\mu_{\tilde{a}}(x) = \begin{cases} \dfrac{x-\underline{a}}{a-\underline{a}} u_{\tilde{a}} & \underline{a} \le x < a \\ u_{\tilde{a}} & x = a \\ \dfrac{\overline{a}-x}{\overline{a}-a} u_{\tilde{a}} & a < x \le \overline{a} \\ 0 & \text{otherwise}, \end{cases} \tag{1}$$

its non-membership function is defined as

$$\vartheta_{\tilde{a}}(x) = \begin{cases} \dfrac{(a-x) + v_{\tilde{a}}(x-\underline{a})}{a-\underline{a}} & \underline{a} \le x < a \\ v_{\tilde{a}} & x = a \\ \dfrac{(x-a) + v_{\tilde{a}}(\overline{a}-x)}{(\overline{a}-a)} & a < x \le \overline{a} \\ 1 & \text{otherwise}, \end{cases} \tag{2}$$

where $u_{\tilde{a}}$ is maximal membership degree, $v_{\tilde{a}}$ is minimal non-membership degree, and $0 \le u_{\tilde{a}} \le 1, 0 \le v_{\tilde{a}} \le 1, 0 \le u_{\tilde{a}} + v_{\tilde{a}} \le 1$.

Let $\pi_{\tilde{a}} = 1 - u_{\tilde{a}} - v_{\tilde{a}}$, $\pi_{\tilde{a}}$ is called the degree of indeterminacy of the element x to $\tilde{a}$, which reflects hesitancy degree of the element x to $\tilde{a}$.

2.2 Einstein Operations

Einstein operations is an effective aggregation operators based on the algebraic t-conorm and t-norm, which includes the Einstein product $\otimes_E$ and the Einstein sum $\oplus_E$ [6]. For any $(a, b) \in [0, 1]^2$, define

$$a \otimes_E b = \frac{a \cdot b}{1 + (1 - a)(1 - b)}, \tag{3}$$

$$a \oplus_E b = \frac{a + b}{1 + ab}. \tag{4}$$

2.3 Einstein Operations of Triangular Intuitionistic Fuzzy Numbers

Definition 2 [23] Let $\tilde{a}_1 = \langle(\underline{a}_1, a_1, \overline{a}_1); u_{\tilde{a}_1}, v_{\tilde{a}_1}\rangle$, $\tilde{a}_2 = \langle(\underline{a}_2, a_2, \overline{a}_2); u_{\tilde{a}_2}, v_{\tilde{a}_2}\rangle$ and $\tilde{a} = \langle(\underline{a}, a, \overline{a}); u_{\tilde{a}}, v_{\tilde{a}}\rangle$ be three triangular intuitionistic fuzzy numbers, and $\lambda \geq 0$, then

(1) $\tilde{a}_1 \otimes_E \tilde{a}_2 = \langle\left(\underline{a}_1\underline{a}_2, a_1a_2, \overline{a}_1\overline{a}_2\right); \frac{u_{\tilde{a}_1}u_{\tilde{a}_2}}{1+(1-u_{\tilde{a}_1})(1-u_{\tilde{a}_2})}, \frac{v_{\tilde{a}_1}+v_{\tilde{a}_2}}{1+v_{\tilde{a}_1}v_{\tilde{a}_2}}\rangle$;

(2) $\tilde{a}^{\wedge_E \lambda} = \langle(\underline{a}^\lambda, a^\lambda, \overline{a}^\lambda); \frac{2u_{\tilde{a}}^\lambda}{(2-u_{\tilde{a}})^\lambda+u_{\tilde{a}}^\lambda}, \frac{(1+v_{\tilde{a}})^\lambda-(1-v_{\tilde{a}})^\lambda}{(1+v_{\tilde{a}})^\lambda+(1-v_{\tilde{a}})^\lambda}\rangle$.

2.4 Ratio Ranking Method of Triangular Intuitionistic Fuzzy Numbers

Let $\tilde{a}_i = \langle(\underline{a}_i, a_i, \bar{a}_i); u_{\tilde{a}_i}, v_{\tilde{a}_i}\rangle (i = 1, 2, \ldots, n)$ be TIFNs. A ratio ranking procedure can be developed for ranking the TIFNs $a_i (i = 1, 2, \ldots, n)$, where $\lambda \in [0, 1]$ is a weight which represents the decision maker's preference information. A ratio ranking process is summarized as follows [7]:

For the same given $\lambda \in [0, 1]$, $i = 1, 2, \ldots, n$, compute

$$R(\tilde{a}_i, \lambda) = \frac{(\underline{a}_i + 4a_i + \overline{a}_i)(\lambda u_{\tilde{a}_i}^2 + (1 - \lambda)(1 - v_{\tilde{a}_i})^2)}{6 + (\overline{a}_i - \underline{a}_i)(\lambda u_{\tilde{a}_i}^2 + (1 - \lambda)(1 - v_{\tilde{a}_i})^2)}, \tag{5}$$

$\lambda \in [0, 1/2]$ shows that the decision maker prefers uncertainty or negative feeling; $\lambda \in [1/2, 1]$ shows that the decision maker prefers certainty or positive feeling; $\lambda = 1/2$ implies that the decision maker is indifferent between positive feeling and negative feeling. In this paper, λ takes 1/2, i.e.,

$$R(\tilde{a}_i) = \frac{(\underline{a}_i + 4a_i + \overline{a}_i)(u_{\tilde{a}_i}^2 + (1 - v_{\tilde{a}_i})^2)}{12 + (\overline{a}_i - \underline{a}_i)(u_{\tilde{a}_i}^2 + (1 - v_{\tilde{a}_i})^2)}. \tag{6}$$

The alternatives can be compared according to non-increasing order of the ratios $R(\tilde{a}_i)(i = 1, 2, \ldots, n)$. The maximum TIFN is the one with the largest ratio.

2.5 Einstein Geometric Aggregation Operators for TIFNs

Definition 3 [23] Let Ω be the set of triangular intuitionistic fuzzy numbers of $\tilde{a}_j = \langle(\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j}\rangle (j = 1, 2, \ldots, n)$. A triangular intuitionistic fuzzy Einstein weighted geometric (TIFEWG) operator of dimension n is mapping $TIFEWG : \Omega^n \rightarrow \Omega$, and

$$TIFEWG(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = \left(\otimes_{j=1}^{n}\right)_E \tilde{a}_j^{\wedge_E \omega_j}. \tag{7}$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weight vector of $\tilde{a}_j (j = 1, 2, \ldots, n)$, with $\omega_j \in [0, 1]$ and $\sum_{j=1}^{n} \omega_j = 1$.

Theorem 1 [23] *Let $\tilde{a}_j = \langle(\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j}\rangle (j = 1, 2, \ldots, n)$ be a collection of triangular intuitionistic fuzzy numbers, then their aggregated value driven from TIFEWG operator is also a triangular intuitionistic fuzzy number, and*

$$\begin{aligned} & TIFEWG(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) \\ & = \left(\otimes_{j=1}^{n}\right)_E \tilde{a}_j^{\wedge_E \omega_j} \\ & = \left\langle \left(\prod_{j=1}^{n} \underline{a}_j^{\omega_j}, \prod_{j=1}^{n} a_j^{\omega_j}, \prod_{j=1}^{n} \overline{a}_j^{\omega_j}\right); \frac{2\prod_{j=1}^{n} u_{\tilde{a}_j}^{\omega_j}}{\prod_{j=1}^{n}(2 - u_{\tilde{a}_j})^{\omega_j} + \prod_{j=1}^{n} u_{\tilde{a}_j}^{\omega_j}}, \right. \\ & \quad \left. \frac{\prod_{j=1}^{n}(1 + v_{\tilde{a}_j})^{\omega_j} - \prod_{j=1}^{n}(1 - v_{\tilde{a}_j})^{\omega_j}}{\prod_{j=1}^{n}(1 + v_{\tilde{a}_j})^{\omega_j} + \prod_{j=1}^{n}(1 - v_{\tilde{a}_j})^{\omega_j}} \right\rangle, \end{aligned} \tag{8}$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weight vector of $\tilde{a}_j (j = 1, 2, \ldots, n)$, with $\omega_j \in [0, 1]$ and $\sum_{j=1}^{n} \omega_j = 1$.

3 Triangular Intuitionistic Fuzzy Einstein Choquet Geometric Operator

In this section, Einstein geometric operator and Choquet integral are employed to aggregate triangular intuitionistic fuzzy numbers given by experts.

Definition 4 [16] A fuzzy measure on X is a set function $\mu : P(X) \rightarrow [0, 1]$, satisfying the following conditions:

(1) $\mu(\phi) = 0, \mu(X) = 1$ (boundary conditions)
(2) If $A, B \in P(X)$ and $A \subseteq B$ then $\mu(A) \leq \mu(B)$ (monotonicity)

$\mu(A \cup B) = \mu(A) + \mu(B) + \rho\mu(A)\mu(B)$ [19], for all $A, B \subseteq P(X)$, and $A \cap B = \phi$, $\rho > -1$.

In particular, if $\rho = 0$, then the condition reduces to the axiom of additive measure: $\mu(A \cup B) = \mu(A) + \mu(B)$ for all $A, B \subseteq X$ and $A \cap B = \phi$ [16].

If $\rho > 0$, then $\mu(A \cup B) > \mu(A) + \mu(B)$, which implies that the set $\{A, B\}$ has multiplicative effect. If $\rho < 0$, then $\mu(A \cup B) < \mu(A) + \mu(B)$, which implies that the set $\{A, B\}$ has substitutive effect [16].

Definition 5 [12] Let f be a positive real-valued function on $X = \{x_1, x_2 \ldots, x_n\}$, and μ be a fuzzy measure on X. The discrete Choquet integral of f with respect to μ is defined by

$$C_\mu(f) = \sum_{i=1}^{n} f_{(i)}[\mu(A_{(i)}) - \mu(A_{(i+1)})], \tag{9}$$

where $A_{(i)} = \{x_{(i)}, \ldots, x_{(n)}\}$, $A_{(n+1)} = \phi$, and $(\cdot)$ indicates a permutation on X such that $f_{(1)} \leq f_{(2)} \leq \cdots \leq f_{(n)}$.

Definition 6 Let $\tilde{a}_j = \langle(\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j}\rangle (j = 1, 2, \ldots, n)$ be a collection of triangular intuitionistic fuzzy numbers on $X = \{x_1, x_2, \ldots, x_n\}$, and μ be a fuzzy measure on X. The discrete intuitionistic fuzzy Choquet integral of $\tilde{a}_j$ with respect to μ is defined as

$$\begin{aligned} &TIFECG_\mu(\tilde{a}_1, \ldots, \tilde{a}_n) \\ &= \tilde{a}_{(1)}^{\wedge_E \mu(A_{(1)}) - \mu(A_{(2)})} \otimes_E \tilde{a}_{(2)}^{\wedge_E \mu(A_{(2)}) - \mu(A_{(3)})} \\ &\quad \otimes_E \cdots \otimes_E \tilde{a}_{(n)}^{\wedge_E \mu(A_{(n)}) - \mu(A_{(n+1)})} \\ &= (\otimes_E)_{j=1}^{n} \tilde{a}_{(j)}^{\wedge_E \mu(A_{(j)}) - \mu(A_{(j+1)})}, \end{aligned} \tag{10}$$

where $(\cdot)$ indicates a permutation on X such that $\tilde{a}_{(1)} \leq \tilde{a}_{(2)} \leq \cdots \leq \tilde{a}_{(n)}$, $A_{(j)} = \{x_{(j)}, \ldots, x_{(n)}\}$, $A_{(n+1)} = \phi$.

Inspired by [23], some desirable properties of the $TIFECG_\mu$ are easy obtained as follows.

Proposition 1 *Let $\tilde{a}_j = \langle(\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j}\rangle (j = 1, 2, \ldots, n)$ be a collection of triangular intuitionistic fuzzy numbers on X, and μ be a fuzzy measure on X. Then their aggregated value driven from $TIFECG_\mu$ operator is also a triangular intuitionistic fuzzy number, and*

$$TIFECG_{\mu}(\tilde{a}_1, \ldots, \tilde{a}_n)$$
$$= \left\langle \left(\Pi_{j=1}^{n} \underline{a}_{(j)}^{\mu(A_{(j)})-\mu(A_{(j+1)})}, \Pi_{j=1}^{n} a_{(j)}^{\mu(A_{(j)})-\mu(A_{(j+1)})}, \Pi_{j=1}^{n} \overline{a}_{(j)}^{\mu(A_{(j)})-\mu(A_{(j+1)})} \right); \right.$$
$$\frac{2\prod_{j=1}^{n} u_{\tilde{a}_{(j)}}^{\mu(A_{(j)})-\mu(A_{(j+1)})}}{\prod_{j=1}^{n} (2 - u_{\tilde{a}_{(j)}})^{\mu(A_{(j)})-\mu(A_{(j+1)})} + \prod_{j=1}^{n} u_{\tilde{a}_{(j)}}^{\mu(A_{(j)})-\mu(A_{(j+1)})}},$$
$$\left. \frac{\prod_{j=1}^{n} (1 + v_{\tilde{a}_{(j)}})^{\mu(A_{(j)})-\mu(A_{(j+1)})} - \prod_{j=1}^{n} (1 - v_{\tilde{a}_{(j)}})^{\mu(A_{(j)})-\mu(A_{(j+1)})}}{\prod_{j=1}^{n} (1 + v_{\tilde{a}_{(j)}})^{\mu(A_{(j)})-\mu(A_{(j+1)})} + \prod_{j=1}^{n} (1 - v_{\tilde{a}_{(j)}})^{\mu(A_{(j)})-\mu(A_{(j+1)})}} \right\rangle, \tag{11}$$

where $(\cdot)$, $A_{(j)}$ $(j = 1, \ldots, n)$ *defined by Definition 6.*

Proposition 2 (Idempotency) *Let* $\tilde{a}_j = \langle (\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j} \rangle (j = 1, 2, \ldots, n)$ *be a collection of triangular intuitionistic fuzzy numbers on X, and* μ *be a fuzzy measure on X. If for any j,* $\tilde{a}_j = \tilde{a} = \langle (\underline{a}, a, \overline{a}); u_{\tilde{a}}, v_{\tilde{a}} \rangle$, *then*

$$TIFECG_{\mu}(\tilde{a}_1, \ldots, \tilde{a}_n) = \tilde{a}.$$

Proposition 3 (Commutativity). *Let* $\tilde{a}_j = \langle (\underline{a}_j, a_j, \overline{a}_j); u_{\tilde{a}_j}, v_{\tilde{a}_j} \rangle$ $(j = 1, 2, \ldots, n)$ *be a collection of triangular intuitionistic fuzzy numbers on X, and* μ *be a fuzzy measure on X. If* $(\tilde{a}_1', \tilde{a}_2', \ldots, \tilde{a}_n')$ *is any permutation of* $(\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n)$, *then* $TIFECG_{\mu}(\tilde{a}_1, \ldots, \tilde{a}_n) = TIFECG_{\mu}(\tilde{a}_1', \tilde{a}_2', \ldots, \tilde{a}_n')$.

4 PROMETHEE II Based on $TIFECG_{\mu}$ Operator

This section constructs the group PROMETHEE II evaluation method, which PROMETHEE II is employed to aggregate evaluation information given by different experts. The decision context includes $A = \{A_1, A_2, \ldots, A_m\}$ be a finite set of alternatives, $C = \{c_1, c_2, \ldots, c_n\}$ be a finite set of attributes, $E = \{E_1, E_2, \ldots, E_K\}$ be the finite set of decision makers, the detail steps are as follows:

Step 1 Identify the weights vector of decision-makers $\mathbf{e} = (e_1, e_2, \ldots, e_K)$, which according to professional knowledge and experience of experts.

Step 2 Construct the decision matrix $D^k = (\tilde{a}_{ij}^k)_{m \times n}$ $(k = 1, 2, \ldots, K)$, in which $\tilde{a}_{ij}^k (i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ is a TIFN provided by E_k for alternative A_i with respect to criterion c_j.

Step 3 Transform the triangular intuitionistic fuzzy decision matrix $D^k = (\tilde{a}_{ij}^k)_{m \times n}$ into the normalized triangular intuitionistic fuzzy decision matrix $R^k = (\tilde{r}_{ij}^k)_{m \times n}$ using the following formulates:

$$\tilde{r}_{ij}^{k} = \begin{cases} \left\langle \left(\frac{\underline{a}_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}}, \frac{a_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}}, \frac{\overline{a}_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}} \right); u_{\tilde{a}_{ij}^{k}}, v_{\tilde{a}_{ij}^{k}} \right\rangle, & j \in Benefit \\ \left\langle \left(1 - \frac{\overline{a}_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}}, 1 - \frac{a_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}}, 1 - \frac{\underline{a}_{ij}^{k}}{(\overline{a}_{j}^{k})^{+}} \right); u_{\tilde{a}_{ij}^{k}}, v_{\tilde{a}_{ij}^{k}} \right\rangle, & j \in Cost \end{cases} \quad (12)$$

where *Benefit* and *Cost* denote the index sets of benefit attributes and cost attributes, respectively, and $(\overline{a}_{j}^{k})^{+} = \max\{\overline{a}_{ij}^{k} | i = 1, 2, \ldots, m\} (j = 1, 2, \ldots, n)$. For convenience, all $\tilde{r}_{ij}^{k} (i = 1, 2, \ldots, m; j = 1, 2, \ldots, n; k = 1, 2, \ldots, K)$ are uniformly denoted by $\tilde{r}_{ij}^{k} = \langle (\underline{r}_{ij}^{k}, r_{ij}^{k}, \overline{r}_{ij}^{k}); u_{r_{ij}^{k}}, v_{r_{ij}^{k}} \rangle$, where $u_{r_{ij}^{k}} = u_{\tilde{a}_{ij}^{k}}, v_{r_{ij}^{k}} = v_{\tilde{a}_{ij}^{k}}$.

Step 4 Determine the fuzzy measures of criteria, and utilize $TIFECG_{\mu}$ operator to aggregate decision information into $\tilde{r}_{i}^{k} = TIFECG_{\mu}(\tilde{r}_{i1}^{k}, \ldots, \tilde{r}_{in}^{k})$.

Step 5 Calculate the preference function $P^{k}(A_i, A_j)$ for the k-th criterion.

$$\text{Gaussian: } P_6(d) = \begin{cases} 1 - e^{(-d^2/2\sigma^2)} & \text{if } d > 0 \\ 0 & \text{if } \ d \le 0 \end{cases} \quad (13)$$

where d denotes the difference between two alternatives with respect to one attribute, σ are parameters [13].

$$P^{k}(d^{k}(A_i, A_j)) = \begin{cases} 1 - e^{(-(d^{k}(A_i, A_j))^2/2\sigma^2)} & \text{if } d^{k}(A_i, A_j) > 0 \\ 0 & \text{if } \ d^{k}(A_i, A_j) \le 0 \end{cases} \quad (14)$$

where $d^{k}(A_i, A_j)$ denotes the difference between ratio values of two alternatives.

$$d^{k}(A_i, A_j) = R(A_i) - R(A_j) \quad (15)$$

and $R(A_i)$ is shown in Eq. (6)

Step 6 Calculate the preference index $\tilde{H}(A_i, A_j)$ as following:

$$\tilde{H}(A_i, A_j) = \sum_{k=1}^{K} e_k P^{k}(A_i, A_j). \quad (16)$$

Step 7 Calculate the positive flow, negative flow and net flow. According to the preference index $\tilde{H}(A_i, A_j)$, the positive flow $\tilde{\phi}^{+}(A_i)$, nagative flow $\tilde{\phi}^{-}(A_i)$ and net flow $\tilde{\phi}(A_i)$ for the i-th alternative can be written as following:

$$\tilde{\phi}^{+}(A_i) = \sum_{j=1}^{m} \tilde{H}(A_i, A_j) \quad (17)$$

$$\tilde{\phi}^{-}(A_i) = \sum_{j=1}^{m} \tilde{H}(A_j, A_i) \quad (18)$$

$$\tilde{\phi}(A_i) = \tilde{\phi}^{+}(A_i) - \tilde{\phi}^{-}(A_i) \quad (19)$$

Step 8 Rank the alternatives. The larger the value of net flow, the better the ranking of $A_i (i = 1, 2, \ldots, m)$.

5 A Numerical Example

Aiming at demonstrating the practicality and effectiveness of the proposed approach, an illustrative example about emergency classification ranking problem is employed, which given by Yue et al. [21]. In this example, three alternatives $A_i (i = 1, 2, 3)$ are ranked by three decision makers E_i $(i = 1, 2, 3)$ referring to three criteria (c_1 scale; c_2 severity degree of loss; c_3 effect to people living).

In the following, the most desirable alternative is selected by applying the proposed approach in Sect. 4.

Step 1 The three DMs' importance weight vector is $\mathbf{e} = (0.35, 0.35, 0.30)^T$.

Step 2 The decision matrixes are showed in Tables 1, 2 and 3 [21].

Step 3 Normalize the TIFN decision matrix. Because the three industrial indices are cost criteria, the normalized TIFN decision matrix is established from Eq. (12), i.e.,

$$R^1 = \begin{pmatrix} \langle(0.50,0.75,0.88);0.70,0.20\rangle & \langle(0,0.40,0.60);0.50,0.40\rangle & \langle(0.22,0.44,0.67);0.70,0.20\rangle \\ \langle(0.25,0.38,0.50);0.60,0.30\rangle & \langle(0,0.20,0.40);0.60,0.30\rangle & \langle(0,0.44,0.56);0.50,0.40\rangle \\ \langle(0,0.50,0.75);0.50,0.40\rangle & \langle(0.20,0.40,0.60);0.80,0.20\rangle & \langle(0.33,0.44,0.89);0.60,0.40\rangle \end{pmatrix}$$

$$R^2 = \begin{pmatrix} \langle(0,0.38,0.62);0.50,0.40\rangle & \langle(0.38,0.62,0.75);0.80,0.20\rangle & \langle(0,0.43,0.71);0.70,0.10\rangle \\ \langle(0.62,0.75,0.88);0.80,0\rangle & \langle(0,0.50,0.62);0.50,0.40\rangle & \langle(0.14,0.43,0.57);0.70,0.20\rangle \\ \langle(0.25,0.62,0.75);0.70,0.20\rangle & \langle(0,0.50,0.88);0.60,0.20\rangle & \langle(0.14,0.57,0.86);0.70,0.20\rangle \end{pmatrix}$$

$$R^3 = \begin{pmatrix} \langle(0,0.25,0.38);0.50,0.20\rangle & \langle(0.12,0.38,0.75);0.60,0.30\rangle & \langle(0.12,0.38,0.75);0.80,0.10\rangle \\ \langle(0.50,0.75,0.88);0.80,0.10\rangle & \langle(0,0.38,0.50);0.70,0.20\rangle & \langle(0.25,0.50,0.88);0.50,0.40\rangle \\ \langle(0.38,0.62,0.75);0.70,0\rangle & \langle(0.12,0.50,0.75);0.50,0.40\rangle & \langle(0,0.38,0.62);0.70,0.20\rangle \end{pmatrix}$$

Step 4 Employ the $TIFECG_\mu$ operator to aggregate decision matrix.
(1) Suppose the fuzzy measures of criteria sets $C = \{c_1, c_2, c_3\}$ are $\mu(c_1) = 0.25$, $\mu(c_2) = 0.37$, $\mu(c_3) = 0.3$, $\mu(c_1, c_2) = 0.65$, $\mu(c_1, c_3) = 0.5$, $\mu(c_2, c_3) = 0.85$, $\mu(c_1, c_2, c_3) = 1$, respectively.
(2) According to Tables 1, 2 and 3, by using the ratio ranking procedure of Eq. (6), the partial evaluation $\tilde{r}_{ij}^k$ of the candidate A_i is reordered such that $\tilde{r}_{i(j)}^k \le \tilde{r}_{i(j+1)}^k (i = 1, 2, 3)(k = 1, 2, 3)$. Then, utilizing the triangular intu-

Table 1 The fuzzy information by decision maker P_1 [21]

Alternatives	c_1	c_2	c_3
A_1	$\langle(1,2,4);0.7,0.2\rangle$	$\langle(2,3,5);0.5,0.4\rangle$	$\langle(3,5,7);0.7,0.2\rangle$
A_2	$\langle(4,5,6);0.6,0.3\rangle$	$\langle(3,4,5);0.6,0.3\rangle$	$\langle(4,5,9);0.5,0.4\rangle$
A_3	$\langle(2,4,8);0.5,0.4\rangle$	$\langle(2,3,4);0.8,0.2\rangle$	$\langle(1,5,6);0.6,0.4\rangle$

Table 2 The fuzzy information by decision maker P_2 [21]

Alternatives	c_1	c_2	c_3
A_1	$\langle(3,5,8);0.5,0.4\rangle$	$\langle(2,3,5);0.8,0.2\rangle$	$\langle(2,4,7);0.7,0.1\rangle$
A_2	$\langle(1,2,3);0.8,0.0\rangle$	$\langle(3,4,8);0.5,0.4\rangle$	$\langle(3,4,6);0.7,0.2\rangle$
A_3	$\langle(2,3,6);0.7,0.2\rangle$	$\langle(1,4,8);0.6,0.2\rangle$	$\langle(1,3,6);0.7,0.2\rangle$

Table 3 The fuzzy information by decision maker P_3 [21]

Alternatives	c_1	c_2	c_3
A_1	$\langle(5,6,8);0.5,0.2\rangle$	$\langle(2,5,7);0.6,0.3\rangle$	$\langle(2,5,7);0.8,0.1\rangle$
A_2	$\langle(1,2,4);0.8,0.1\rangle$	$\langle(4,5,8);0.7,0.2\rangle$	$\langle(1,4,6);0.5,0.4\rangle$
A_3	$\langle(2,3,5);0.7,0.0\rangle$	$\langle(2,4,7);0.5,0.4\rangle$	$\langle(3,5,8);0.7,0.2\rangle$

itionistic fuzzy Einstein Choquet to derive the individual overall preference triangular intuitionistic fuzzy values $\tilde{r}_i^k$ of the alternative A_i.

$$\begin{aligned}
&\tilde{r}_1^1 = \langle(0,0.48,0.68);0.60,0.30\rangle,\ \tilde{r}_2^1 = \langle(0,0.29,0.46);0.57,0.33\rangle,\\
&\tilde{r}_3^1 = \langle(0,0.43,0.75);0.65,0.33\rangle,\ \tilde{r}_1^2 = \langle(0,0.48,0.71);0.70,0.18\rangle,\\
&\tilde{r}_2^2 = \langle(0,0.53,0.66);0.62,0.26\rangle,\ \tilde{r}_3^2 = \langle(0,0.55,0.84);0.65,0.20\rangle,\\
&\tilde{r}_1^3 = \langle(0,0.36,0.68);0.64,0.23\rangle,\ \tilde{r}_2^3 = \langle(0,0.50,0.70);0.65,0.25\rangle,\\
&\tilde{r}_3^3 = \langle(0,0.49,0.72);0.60,0.26\rangle.
\end{aligned}$$

Step 5 Construct the priority function.
Combined the difference of the TIFNs' ratio with the Gaussian function, the preference functions for the pairs of alternatives are computed from Eqs. (6) and (14), where σ takes 0.25, then the results are shown in Table 4.

Step 6 Calculate the preference index $\tilde{H}(A_i, A_j)$.

$$\begin{aligned}
&\tilde{H}(A_1,A_2) = 0.30, \tilde{H}(A_1,A_3) = 0.00\\
&\tilde{H}(A_2,A_1) = 0.09, \tilde{H}(A_2,A_3) = 0.02\\
&\tilde{H}(A_3,A_1) = 0.04, \tilde{H}(A_3,A_2) = 0.33.
\end{aligned}$$

Step 7 Calculate fuzzy positive flow, negative flow and net flow. The fuzzy positive flow, negative flow and net flow are obtained in Table 5.

Table 4 Preference functions for the pairs of alternatives

i	1	2	3
$P_i(A_1, A_2)$	0.67	0.20	0
$P_i(A_1, A_3)$	0.01	0	0
$P_i(A_2, A_1)$	0	0	0.30
$P_i(A_2, A_3)$	0	0	0.06
$P_i(A_3, A_1)$	0	0.03	0.11
$P_i(A_3, A_2)$	0.61	0.33	0

Table 5 The ranking of alternatives

Alternatives	$\phi^+(i)$	$\phi^-(i)$	$\phi(i)$	Ranking
A_1	0.30	0.13	0.17	2
A_2	0.11	0.63	−0.52	3
A_3	0.37	0.02	0.35	1

Step 8 Rank all the alternatives $A_i(i = 1, 2, 3)$ in accordance with the overall preference values.

$$A_3 \succ A_1 \succ A_2$$

Thus the most desirable alternative is A_3.

The ranking result obtained by the proposed PROMETHEE II ranking method is different from those obtained by Yue et al. [21], which driven from different aggregation operators. In real life, decision makers can properly choose the desirable alternative according to his actual needs by adjusting interaction grade.

6 Conclusion

In this paper, a new group PROMETHEE II method is established to solve the MAGDM problems, in which the ratings of alternatives on attributes are expressed with TIFNs and triangular intuitionistic fuzzy Einstein Choquet integral operator to reflect the interactions phenomena among the decision making criteria. An example of emergency classification is included to illustrate the method which is more feasible and practical for real decision making problems. The future work of this paper is to improve the aggregated algorithm and investigate the operators to many other practical MAGDM problems, such as decision making, patter recognition and clustering analysis and so on.

Acknowledgments Thanks to the supported by the Natural Science Foundation of China (No. 61203283), Liaoning Provincial Natural Science Foundation of China (No. 2014025004) and the

Fundamental Research Funds for the Central Universities (Nos. 3132014036, 3132014324), Scientific Research Project of Liaoning Provincial Education Department (No. L2015072).

References

1. Beg, I., Rashid, T.: Multi-criteria trapezoidal valued intuitionistic fuzzy decision making with Choquet integral based TOPSIS. Opsearch **51**(1), 98–129 (2014)
2. Chuu, S.J.: Evaluating the flexibility in a manufacturing system using fuzzy multi-attribute group decision-making with multi-granularity linguistic information. Int. J. Adv. Manuf. Technol. **32**(3–4), 409–421 (2007)
3. Grabisch, M.: Fuzzy integral in multicriteria decision making. Fuzzy Sets Syst. **69**(3), 279–298 (1995)
4. Grabisch, M., Sugeno, M., Murofushi, T.: Fuzzy measures and integrals: theory and applications. Springer-Verlag New York, Inc. (2000)
5. Ju, Y., Yang, S.: Approaches for multi-attribute group decision making based on intuitionistic trapezoid fuzzy linguistic power aggregation operators. J. Intell. Fuzzy Syst. **27**(2), 987–1000 (2014)
6. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Position paper I: basic analytical and algebraic properties. Fuzzy Sets Syst. **143**(1), 5–26 (2004)
7. Li, D.F.: A ratio ranking method of triangular intuitionistic fuzzy numbers and its application to MADM problems. Comput. Math. Appl. **60**(6), 1557–1570 (2010)
8. Liu, X., Ju, Y., Wang, A.: A multiple attribute group decision making method with its application to emergency alternative assessment. J. Converg. Inf. Technol. **7**(2) (2012)
9. Meyer, P., Roubens, M.: On the use of the Choquet integral with fuzzy numbers in multiple criteria decision support. Fuzzy Sets Syst. **157**(7), 927–938 (2006)
10. Mishra, S., Samantra, C., Datta, S., Mahapatra, S.S.: Multi-attribute group decision-making (MAGDM) for supplier selection using fuzzy linguistic modelling integrated with VIKOR method. Int. J. Serv. Oper. Manage. **12**(1), 67–89 (2012)
11. Shu, M.H., Cheng, C.H., Chang, J.R.: Using intuitionistic fuzzy sets for fault-tree analysis on printed circuit board assembly. Microelectron. Reliab. **46**(12), 2139–2148 (2006)
12. Tan, C.X., Chen, X.H.: Intuitionistic fuzzy Choquet integral operator for multi-criteria decision making. Expert Syst. Appl. **37**(1), 149–157 (2010)
13. Vincke, J.P., Brans, P.: A preference ranking organization method. The PROMETHEE method for MCDM. Manage. Sci. **31**(6), 647–656 (1985)
14. Wang, J.Q., Zhong, Z.: Aggregation operators on intuitionistic trapezoidal fuzzy number and its application to multi-criteria decision making problems. J. Syst. Eng. Electron. **20**(2), 321–326 (2009)
15. Wan, S.P.: Power average operators of trapezoidal intuitionistic fuzzy numbers and application to multi-attribute group decision making. Appl. Math. Model. **37**(6), 4112–4126 (2013)
16. Wang, Z.Y., Klir, G.J.: Fuzzy Measure Theory. Springer Science & Business Media, New York (1992)
17. Wei, G.W., Lin, R., Zhao, X.F., Wang, H.J.: An approach to multiple attribute decision making based on the induced Choquet integral with fuzzy number intuitionistic fuzzy information. J. Bus. Econ. Manage. **15**(2), 277–298 (2014)
18. Wu, J., Liu, Y.: An approach for multiple attribute group decision making problems with interval-valued intuitionistic trapezoidal fuzzy numbers. Comput. Ind. Eng. **66**(2), 311–324 (2013)
19. Xu, Y., Wang, H., Merig, J.M.: Intuitionistic fuzzy Einstein Choquet integral operators for multiple attribute decision making. Technol. Econ. Dev. Econ. **20**(2), 227–253 (2014)
20. Yu, D.: Intuitionistic fuzzy Choquet aggregation operator based on Einstein operation laws. Sci. Iran. Trans. E Ind. Eng. **20**(6), 2109 (2013)

21. Yue, X., Xia, G.K., Li, Y.: Multi-attribute group decision-making method based on triangular intuitionistic fuzzy number and 2-tuple linguistic information. J. Softw. **7**(7), 1546–1553 (2012)
22. Zhang, S., Yu, D.: Some geometric Choquet aggregation operators using Einstein operations under intuitionistic fuzzy environment. J. Intell. Fuzzy Syst.: Appl. Eng. Technol. **26**(1), 491–500 (2014)
23. Zhao, S.P., Liang, C.Y., Zhang, J.L.: Some intuitionistic trapezoidal fuzzy aggregation operators based on Einstein operations and their application in multiple attribute group decision making. Int. J. Mach. Learn. Cybern. (2015). doi:10.1007/s13042-015-0349-2

Dynamics Analysis and Fuzzy Control for the Working Device of Hydraulic Excavator

Lv Guang-ming, Xu Guo-sheng and Shuai Xie

Abstract The Kane-Huston is adopted in this paper to build the topological structure for the working device of hydraulic excavator. And the virtual prototyping model of the working device is built by the virtual prototype analysis software. Through analysis of the mechanical properties of working device on real working condition and comparing the inverse dynamics gotten by the Kane-Huston with that gotten by the virtual prototype model in software, the correctness of the dynamic equations from the Kane-Huston is verified. Finally, fuzzy control method is used for the optimization of bucket tip trajectory and it has good control effect in the software simulation environment.

Keywords Hydraulic excavator · Working device · Dynamics · Trajectory · Fuzzy control

1 Introduction

At present, Newton-Euler and Lagrange equation [1, 2] are widely used in the dynamics research. For the complex logical deduction and hard solution in these equations, the Huston method based on Kane equation [3] is used in this paper to analyse the working device of hydraulic excavator which involves boom, arm and bucket. As [4] reports, Huston method based on Kane equation can effectively solve

L. Guang-ming (✉)
School of Mechatronics Engineering, Harbin Institute of Technology,
Harbin 150001, Nan Gang District, China
e-mail: 1043612258@qq.com; lgmhit@hit.edu.cn

X. Guo-sheng · S. Xie
School of Mechatronics Engineering, Harbin Institute of Technology,
Harbin 150001, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_16

the problem of multi-body dynamics with low-order body arrays to describe the system's topology structure.

After dynamic modeling of working device with Kane-Huston method, it is necessary to verify the correctness of the established dynamics equation, thus ADAMS is used to build a virtual prototype model of the working device of the hydraulic excavator. And through the analysis of the results in inverse dynamics respectively from the dynamics equation based on Kane-Huston and the virtual prototype model in ADAMS, the correctness of the theoretical dynamics model based on Kane-Huston is verified. The virtual prototype model of the working device under the simulated digging environment is analysed, and the mechanical properties of the linkage mechanism which is composed of boom, arm and bucket are obtained, providing the basis for the analysis of strength check and control strategy.

Finally fuzzy control is used in the bucket tip trajectory optimization. Though detecting the change of the boom angle and arm angle, fuzzy unit can improve the accuracy in trajectory control of the bucket tip under ADAMS and MATLAB [5] co-simulation.

2 Dynamic Modeling of the Working Device Based on Kane-Huston Equations

Hydraulic excavator working device is mainly composed of boom, boom cylinder, arm, arm cylinder, bucket and bucket cylinder. It can be seen as the plane open kinematic chain with four degrees of freedom despite consideration of rotating motion and walking movement. Three cylinders can be simplified as driving torque M_1, M_2 and M_3, changing over time. The simplified model is shown in Fig. 1.

There, the boom is marked as B_1, the length of it is l_1 and its center of mass to the hinge joint A is r_1; the length of the arm marked as B_2 is l_2, its center of mass to the hinge joint B is r_2; the length of the bucket marked as B_3 is l_3, its center of mass to the hinge joint C is r_3. The external forces include the torque of M_1, M_2, M_3 respectively acting on B_1, B_2 and B_3, and the gravity m_1g, m_2g, m_3g, respectively of B_1, B_2 and B_3. The fixed reference frame is $An_1^1n_2^1$, Angle between the rotation platform and B_1 is θ_1, and it's positive when B_1 moves counterclockwise relative to the fixed rotation platform, the driving torque on B_1 is M_1. The $An_1^2n_2^2$, $Bn_1^3n_2^3$ and $Cn_1^4n_2^4$ are the generalized coordinate frames. Angle between B_2 and B_1 is θ_2, and it's positive when B_2 moves clockwise relative to the B_1, the driving torque on B_2 is M_2. Angle between B_3 and B_2 is θ_3, and it's positive when B_3 moves clockwise relative to the B_2, the driving torque on B_3 is M_3.

If $(\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3) = (u_1, u_2, u_3)$, according to Kane-Huston method, the relationship between driving torque M_1, M_2, M_3 respectively working on the boom, arm, bucket and the angle θ_1, θ_2, θ_3, can be presented as follows.

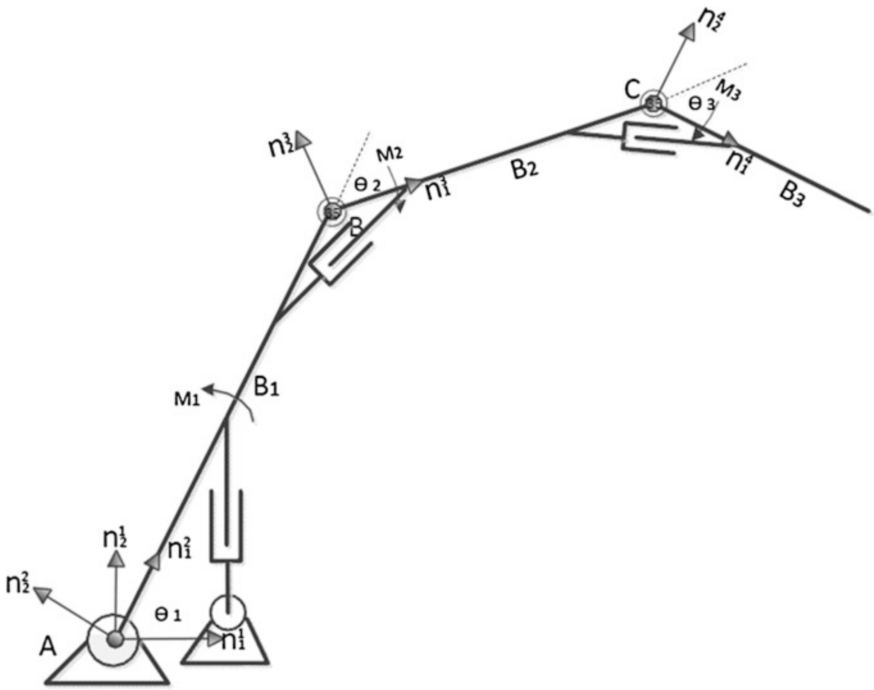


Fig. 1 Simplified model for the working device

$$
\begin{aligned}
&(-m_1 r_1^2 - m_2 l_1^2 - m_2 l_1 r_2 c_2 - m_3 l_1^2 - m_3 l_1 l_2 c_2 - m_3 l_1 r_3 c_{2+3} + 2m_3 l_2 r_3 c_3 + 2m_3 r_3^2 - I_1 + 2I_3)\,\dot{u}_1 \\
&+ (m_2 l_1 r_2 + m_3 l_1 l_2 c_2 + m_3 l_1 r_3 c_{2+3} - 2m_3 l_2 r_3 c_3 - 2m_3 r_3^2 - 2I_3)\,\dot{u}_2 + (m_3 l_1 r_3 c_{2+3} - 2m_3 r_3^2 - 2I_3)\,\dot{u}_3 \\
&+ (-m_3 l_1 r_3 s_{2+3} - m_3 l_2 r_3 c_3 + m_3 l_2 r_3 s_3)(u_1 - u_2 - u_3)^2 + (-m_3 l_1 l_2 s_2 - m_3 l_2 r_3 c_3 - 3m_3 l_2 r_2 s_3)(u_1 - u_2)^2 \\
&+ (-m_3 l_1 l_2 s_3 - m_2 l_1 r_2 s_2 - m_3 l_1 l_2 s_2 - 4m_3 l_1 r_3 s_{2+3})u_1^2 - m_2 u_2^2 + 2m_2 l_1 r_2 s_2 u_1 u_2 - m_1 g r_1 c_1 - m_2 g l_1 c_1 \\
&- m_2 g r_2 c_{1-2} - m_3 g l_1 c_1 - m_3 g r_3 c_{1-2-3} - 2m_3 g l_2 c_{1-2} - m_2 g r_3 c_{1-2} - 3m_2 g r_3 c_{1-2-3} + M_1 = 0
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
&(m_2 l_1 r_2 c_2 + m_3 l_2 r_3 c_3 + m_2 r_2^2 + m_3 l_1 l_2 c_2 + m_3 l_2^2 + 2m_3 l_2 r_3 c_3 + 2m_3 l_1 r_3 c_{2+3} + 2m_3 r_3^2 + I_2 + 2I_3)\,\dot{u}_1 \\
&- (m_2 r_2^2 + m_3 l_2^2 + 3m_3 l_2 r_3 c_3 + 2m_3 r_3^2 + I_2 + 2I_3)\,\dot{u}_2 - (2m_3 r_3^2 + m_3 l_2 r_3 c_3 + 2I_3)\,\dot{u}_3 - (m_2 l_1 r_2 s_2 \\
&+ m_3 l_2 r_3 s_3 + 2m_3 l_1 r_3 s_{2+3} + m_3 l_1 l_2 s_2)u_1^2 + 2m_3 l_2 r_3 s_3 u_1 u_2 - 2m_3 l_2 r_3 s_3 u_1 u_3 + m_3 l_2 r_3 s_3 u_3^2 + 2m_3 l_2 r_3 s_3 u_2 u_3 \\
&\quad - m_3 l_2 r_3 s_3 u_2^2 - m_3 g r_3 c_{1-2-3} - m_3 g(l_2 c_{1-2} + r_3 c_{1-2-3}) - m_2 g r_3 c_{1-2} + M_2 = 0
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
&(m_3 l_1 r_3 c_{2+3} + m_3 l_2 r_3 c_3 + m_3 r_3^2 + I_3)\,\dot{u}_1 - (m_3 l_2 r_3 c_3 + m_3 r_3^2 + I_3)\,\dot{u}_2 - (m_3 r_3^2 + I_3)\,\dot{u}_3 \\
&- (m_3 l_1 r_3 s_{2+3} + m_3 l_2 r_3 s_3)u_1^2 + 2m_3 l_2 r_3 s_3 u_1 u_2 - m_3 l_2 r_3 s_3 u_2^2 - m_3 g r_3 c_{1-2-3} + M_3 = 0
\end{aligned} \tag{3}
$$

There, I_1, I_2, I_3 are the rotational inertia of the B_1, B_2, B_3 respectively relative to the hinge joints A, B, C. Other symbols that are as follows to illustrate.

$$\begin{array}{lll} c_1 = \cos\theta_1, & c_{1-2} = \cos(\theta_1 - \theta_2), & c_{1-2-3} = \cos(\theta_1 - \theta_2 - \theta_3) \\ s_1 = \sin\theta_1, & s_{1-2} = \sin(\theta_1 - \theta_2), & s_{1-2-3} = \sin(\theta_1 - \theta_2 - \theta_3) \end{array} \tag{4}$$

3 Virtual Prototype Simulation of Working Device Based on the ADAMS

3.1 Virtual Prototype Simulation

The driving functions in ADAMS of each hydraulic cylinder and revolute pair are given based on the actual mining of the working device and the theoretical cycle schedule of operation. Time of the cycle period is 16.2 s. And it realizes the movement of the digging, elevating, unloading and arm dropping, etc. The function of tangential resistance W1 in ADAMS is:

STEP(time, 1.0, 0.0, 4.0, 22961) + STEP(time, 4.0, 0.0, 5.0, −22961)
The function of the normal resistance W2 in ADAMS is:
STEP(time, 1.0, 0.0, 4.0, 9644.0) + STEP(time, 4.0, 0.0, 5.0, −9644)
The function of the the material weight G in ADAMS is:
STEP(time, 1.0, 0.0, 5.0, 4939.2) + STEP(time, 5.0, 0.0, 10.3, 0) + STEP(time, 10.3, 0.0, 13.1, −4939.2)

The simulation results are shown in Fig. 2.

3.2 Inverse Dynamic Problem from Virtual Prototype Simulation and Kane-Huston Equation

First, in ADAMS simulation environment, the angle between bucket and arm is locked as 30°, and then the model movement of the working device is conducted on a straight trajectory from (4000, 0) to (7000, 0) in the coordinate frame. So the Angle change curves of boom and arm is obtained. MATLAB is used to solve the inverse dynamic equation based on the Kane-Huston and also get the angle change curves of the boom and arm.

The results are compared as follows.

From Figs. 3 and 4, it can be concluded that the angle change curves of the boom and arm source from the ADAMS simulation and the Kane-Huston dynamics equation are basically consistent within 0.5° Angle deviation. Model is well established.

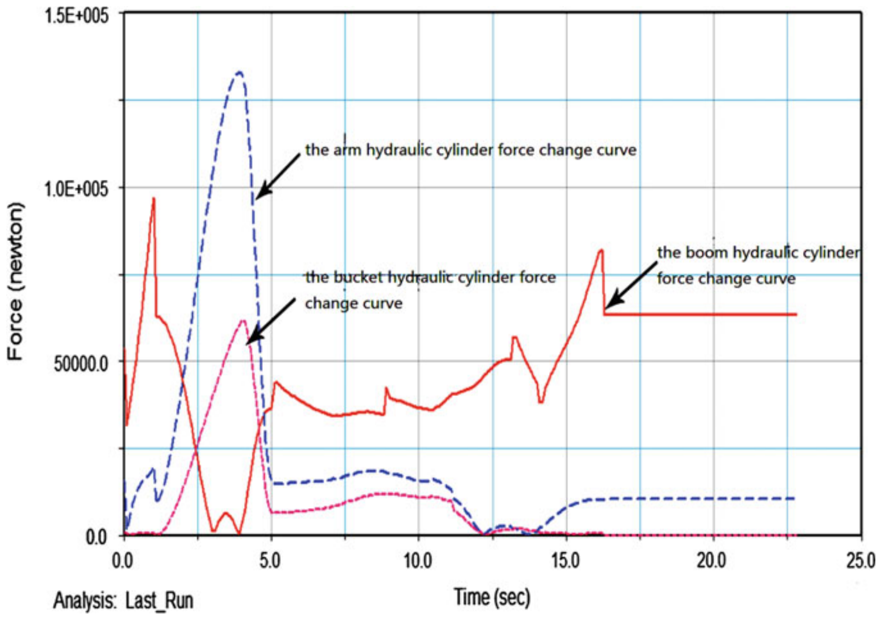


Fig. 2 Forces change curves of boom, arm and bucket hydraulic cylinders

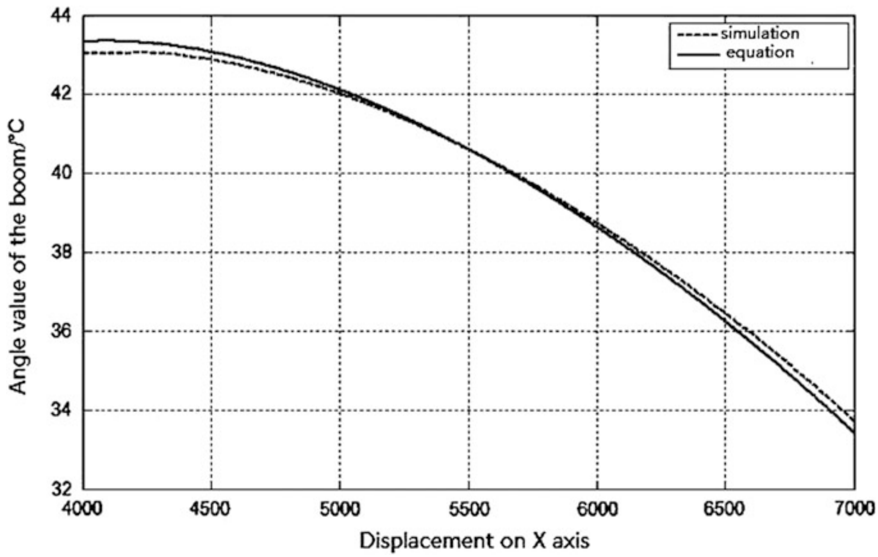


Fig. 3 Angle change curves of the boom source from ADAMS simulation and Kane-Huston equation

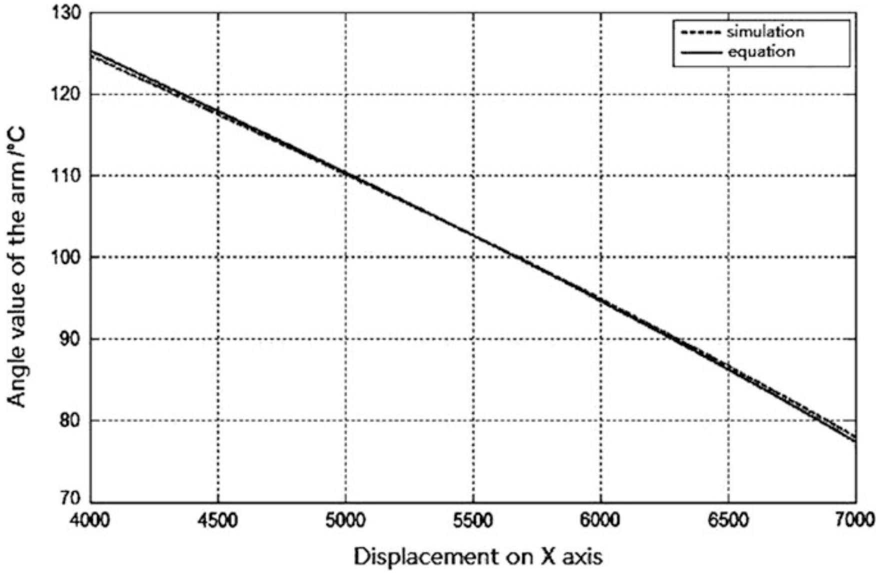


Fig. 4 Angle change curves of the arm source from ADAMS simulation and Kane-Huston equation

4 The Fuzzy Control Simulation of Virtual Prototype of Working Device

The fuzzy control simulation of virtual prototype of the working device is under ADAMS and MATLAB/SIMULINK co-simulation. First the 3D model of the working device is added the constraints and drive functions which are the analog of the real working environment of the hydraulic excavator in ADAMS, and then the related data and parameters describing the system are sent from ADAMS to MATLAB/SIMULINK. And finally, the fuzzy control plan is established in the SIMULINK. The block diagram of virtual prototype control system is shown in Fig. 5.

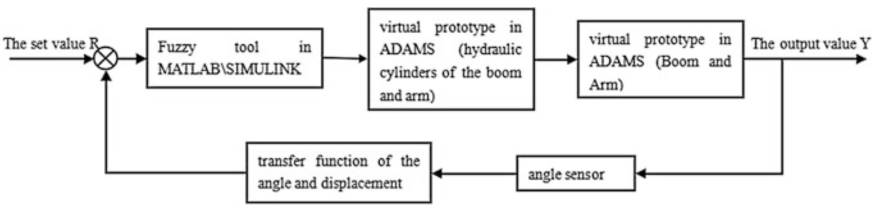


Fig. 5 The block diagram of control system

Figure 6 is shown the fuzzy control system of the working device. DBzhuangtaishuru and DGzhuangtaishuru in Fig. 7 are respectively the input driving functions of the hydraulic cylinders of the boom and arm derived from MATLAB/SIMULINK. DBzhuangtaishuchu and DGzhuangtaishuchu in Fig. 7 are

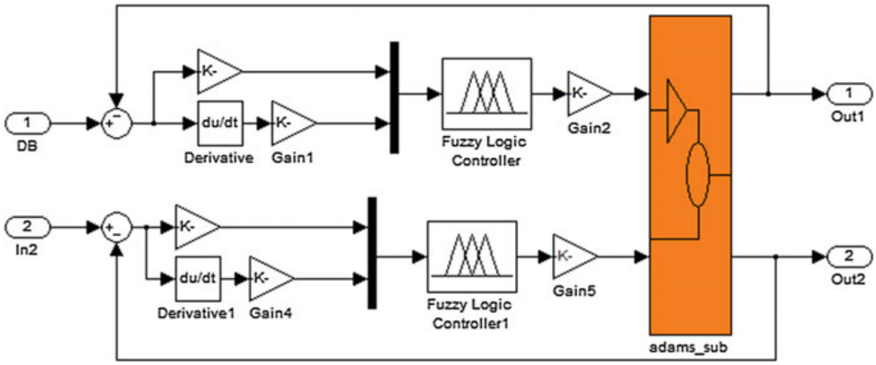


Fig. 6 Module connection of fuzzy control

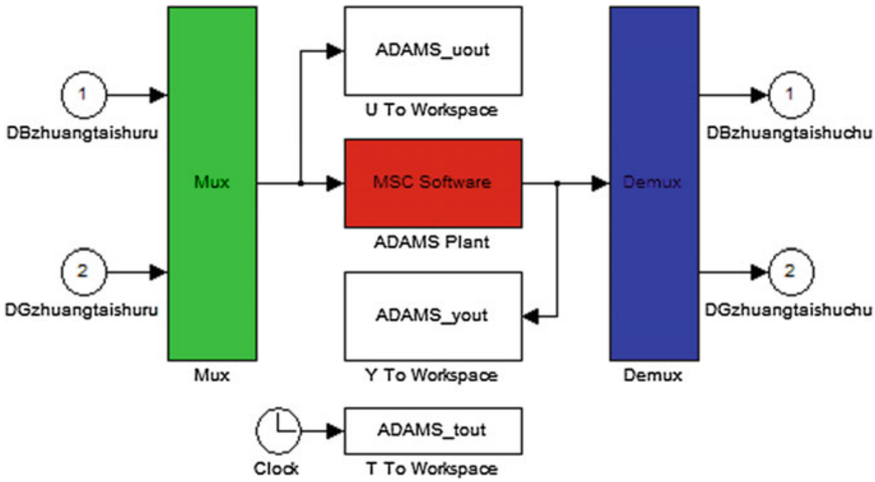


Fig. 7 The data exchange of the block 'adams_sub'

respectively the output angle change of the boom and arm in ADAMS. The input of the fuzzy inference for switching is error(E) and the differential of the error(EC), the output is coefficient(U) of the fuzzy controller.

Figures 8 and 9 show the membership functions of the E, EC and U. Table 1 shows the control rules.

Finally the results of the bucket tip trajectory are gotten as Figs. 10 and 11.

From Figs. 10 and 11, it can be concluded that the bucket tip trajectory is much more consistent with the designed trajectory using fuzzy control.

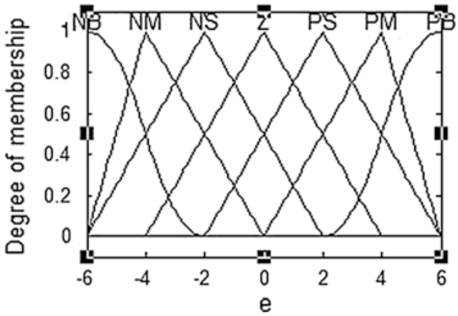


Fig. 8 The membership function for E or EC

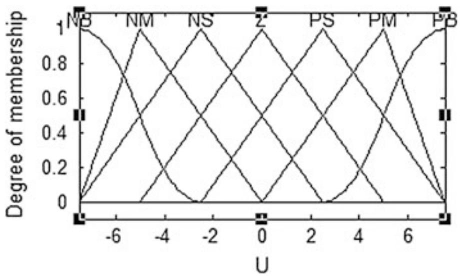


Fig. 9 The membership function for U

Table 1 Fuzzy control rules

E \ EC (U)	PB	PM	PS	Z	NS	NM	NB
PB	NB	NB	NM	NM	NS	NS	Z
PM	NB	NM	NM	NS	NS	Z	PS
PS	NM	NM	NS	NS	Z	PS	PS
Z	NM	NS	NS	Z	PS	PS	PM
NS	NS	NS	Z	PS	PS	PM	PM
NM	NS	Z	PS	PS	PM	PM	PB
NB	Z	PS	PS	PM	PM	PB	PB

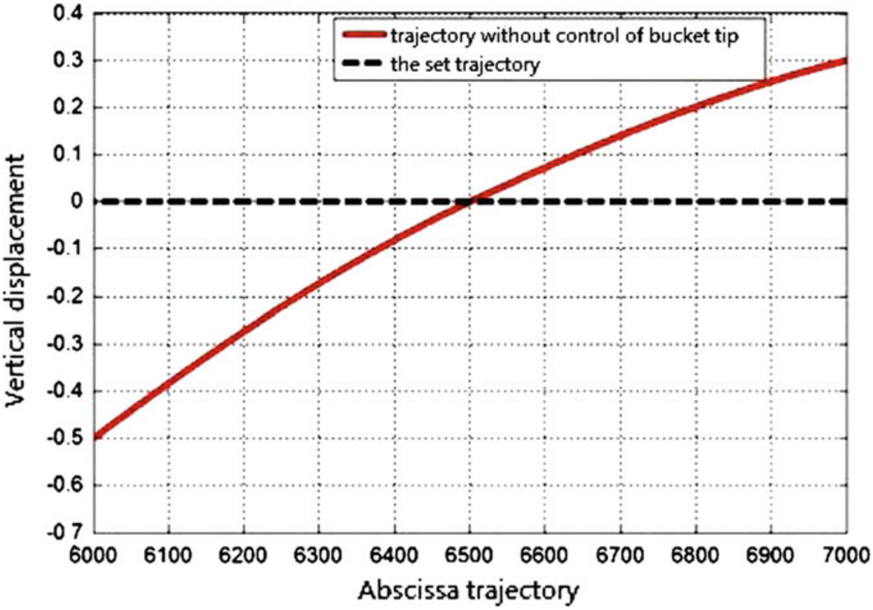


Fig. 10 The trajectory without Fuzzy control

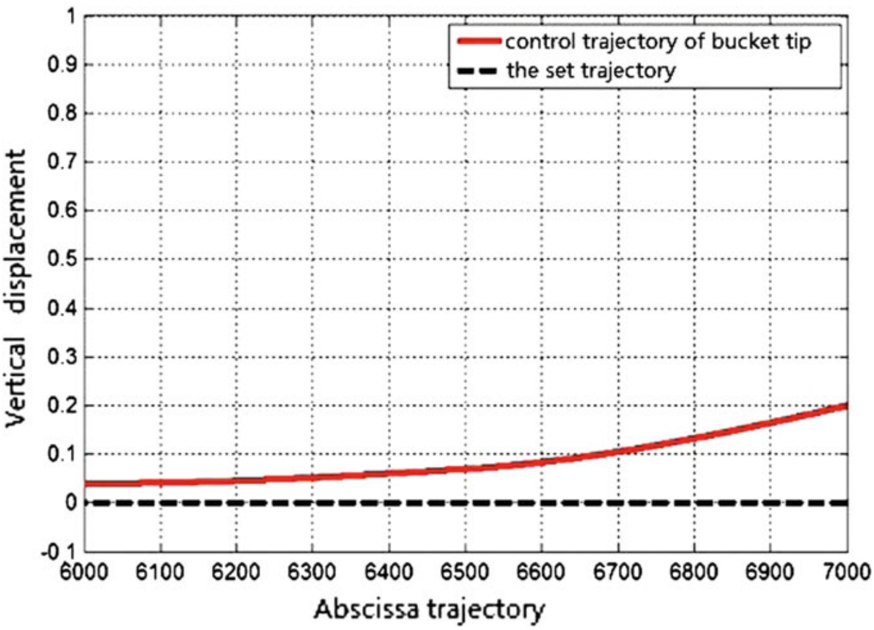


Fig. 11 The trajectory with Fuzzy control

5 Conclusion

In the paper, the model of the working device of hydraulic excavator is built based on the Kane-Huston method in multi-body dynamics. And through comparing angle change curves of the boom and arm from solving the inverse dynamic equation based on the Kane-Huston in MATLAB with the angle change curves of the boom and arm gotten by ADAMS, the dynamics equations based on Kane-Huston are verified correct.

Finally fuzzy control is used in the bucket tip trajectory control and it has a good control effect in the ADAMS and MATLAB co-simulation environment.

References

1. Kim, Y.B., Ha, J., Kang, H., Kim, P.Y., Park, J., Park, F.C.: Dynamically optimal trajectories for earthmoving excavators. J. Autom. Constr. **35**, 568–578 (2013)
2. Haga, M., Hiroshi, W., Fujishima, K.: Digging control system for hydraulic excavator. J. Mechatron. **11**(6), 665–676 (2001)
3. Šalinić, S., Bošković, G., Nikolić, M.: Dynamic modeling of hydraulic excavator motion using Kane's equations. J. Autom. Constr. **44**, 56–62 (2014)
4. Deng, Z., Wang, L., Zhang, L., Fu, Y.: The establishment and dynamic simulation of the virtual prototype of a single bucket backhoe hydraulic excavator's working device. J. Mach. Electron (12), 33–35 (2009) (in Chinese)
5. MSC. Software Corp. Using the ADAMS/View Function Builder. MSC. Software Corporation, USA (2003)

Part II
Fuzzy Engineering

The Application and Predictive Models Base on Bayesian Classifier in Electronic Information Industry

Yu-Bin Zhong, Tao Li, Si-cong Chen and Si-cong Guo

Abstract This article analyzes the 120 months level of electronic information industry development data of 35 indicators in 30 provinces from the ZhongHong database. A Bayesian Classifier forecast model is built to measure the development level of electronic information industry index α, considering four core modeling properties which are smoothness, randomness, trends and normality. This model is used for simulation and forecasts of actual application of the electronic information industry. Specifically, according to the variation of index α in each region at different time, time-series forecast model is built for the forecasts of index α in each region in 2014; Bayesian classifier model is then constructed to evaluate the development level of electronic industry given the index α in each region. The Bayesian classifier can classify different α into year so that it can produce a development trend of electronic industry in each province. The results indicate that the development levels of electronic industry in Shanghai and Beijing is higher than other provinces. However, the differences between the levels in the two advanced cities and Zhejiang, Tianjin, Jiangsu, Fujian and Guangdong are shrinking from 2013 to 2014. In addition, the results also show that there are no record in the evaluation of 'slightly advanced' region. However, the electronic industry in Shandong, Hubei, Liaoning, Ningxia and Hainan will have good prospects in 2014, filling up the area of 'slightly advanced' region. This model requires not much data, which can be easily transplanted. The model is suitable for research and analysis of macro mechanism without much data and is also a reference for government department and related businesses.

Y.-B. Zhong (✉) · T. Li
School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China
e-mail: zhong_yb@163.com

S. Chen
School of Civil Engineering, Guangzhou University, 510006 Guangzhou, China

S. Guo
Institute of Intelligence and Mathematics, Liaoning Technical University, Fuxin 123000, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_17

Keywords Electronic information industry · Bayesian classifier · Time series forecasting · Industry development index

1 Introduction

The achievement of industrialization and informatization is the strategic target of economic and social development of our country. With the rapid growth of China's economy, the level of Chinese information technology has a significant development and progress, the current has been to the Internet and cloud computing as the main representative of emerging social productivity forward. How to evaluate the level of information development of major provinces and municipalities in China is an important issue at present.

This project intends to achieve the following two tasks:

(1) According to the known data, the electronic information industry development level of our country is evaluated by using Bayesian classifier;
(2) A reasonable forecast of the development level of the electronic information industry in the provinces of the future for a period of time.

In January 2014, China's Ministry of industry and information technology combined with China's Electronic Information Industry Development Research Institute released the 2013 the development level of China's information technology assessment report (hereinafter referred to as the "report"), the "report" with complete index system of China's 31 provinces, cities and autonomous regions the level of information to make the score. In this project, the parameter range of the Bayesian classifier is defined by the score level of the report [1].

The project data sources in Zhonghong database, a collection of the report contained in the national 30 major provinces, cities and autonomous regions of the

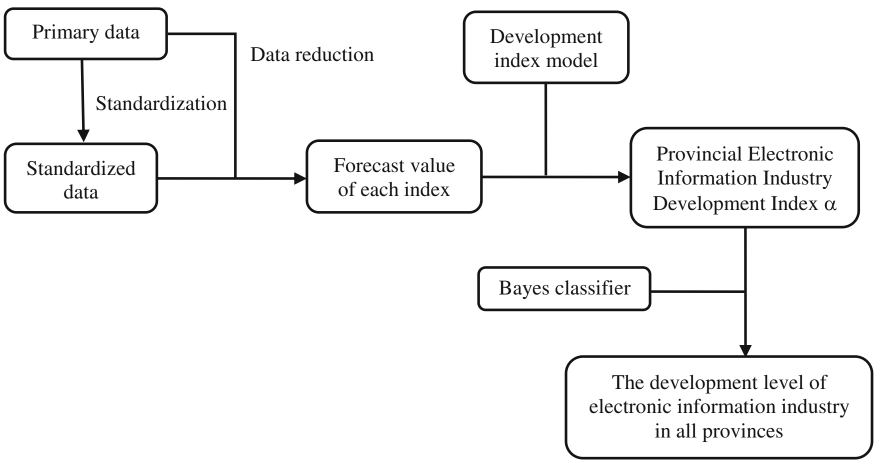


Fig. 1 Technical roadmap for research

electronic information industry of 35 indicators of data, the time span of the data of each index for June 2003 and October 2013, the data time window width for 1 month.

This project is carried out in accordance with the following technical route (Fig. 1).

2 Preparation for Modeling

2.1 Assumed Condition

(1) The assumption that the stability of the same province in a period of time to maintain the stability of the timing. The hypothesis that taking into account the economic and social actual operation situation, industry development has a certain inertia; at the same time the industrial policies with a lag, so the index data will gradual and not mutation. The hypothesis is consistent with the observed data trends, and is also the basis for interpolation of missing data;
(2) The assumption that every index of every province can be regarded as the observation value of a sequence of random processes [2], which is the basis of the use of the Bayesian classifier and the establishment of the prediction model;
(3) Trend assumption: each one class index was normalized after the results can be used with a timing model to predict, the hypothesis depicts the economic activity of each endogenous variable between have a strong linear relationship between the background [3], based on the assumption that we only need to time series prediction model to do a test can reduces the algorithm's time and space complexity;
(4) Normal assumption [4]: if there is no time for each indicator to meet the normal distribution of a parameter. Note that this hypothesis does not require the same distribution, and there is no contradiction with the assumption of stationary, because the stationary is only a statistical significance of the observed results. The hypothesis is used to simplify the parameters training of Bayesian classifier.

2.2 Index System and Data Preprocessing

2.2.1 Constitute the Index System

According to the data provided by the macro database, select 35 indicators to construct the evaluation index system, as shown in Table 1.

Table 1 Evaluation index system of USRC elements

Index category	Specific index	Unit
Total quantity index (this year per month)	Cumulative number of enterprises	Individual
	Cumulative loss of business units	Individual
	The cumulative losses of the enterprise	Thousand yuan
	Total loss of business losses than the same period last year growth rate	%
Assets index (this year per month)	The cumulative net accounts receivable	Thousand yuan
	The cumulative net receivables growth over the same period last year	%
	Finished product	Thousand yuan
	Cumulative production over the same period last year	%
	Average balance of current assets	Thousand yuan
	Average balance of current assets over the same period last year growth	%
	Accumulated assets	Thousand yuan
	Accumulated assets over the same period last year growth	%
	Total liabilities for this month	Thousand yuan
	Total liabilities increased over the same period last year	%
Cost index (this year per month)	Product sales revenue	Thousand yuan
	Cumulative product sales revenue over the same period last year	%
	Cumulative product sales cost	Thousand yuan
	Cumulative product sales increased over the same period last year	%
	Cumulative product sales	Thousand yuan
	Cumulative product sales expenses increased over the same period last year	%
	Accumulated product sales tax and additional	Thousand yuan
	Cumulative product sales tax and the growth of the same period last year	%
	Accumulated financial expenses	Thousand yuan
	Cumulative financial expenses increased over the same period last year	%
	Accumulated management cost	Thousand yuan
	Cumulative management costs over the same period last year	%
	Total profit	Thousand yuan
	Total profit growth over the same period last year	%

(continued)

Table 1 (continued)

Index category	Specific index	Unit
Growth index	Cumulative total number of employees (this year per month)	People
	Cumulative total number of employees increased over the same period last year(this year per month)	%
	Total tax amount (this year per month)	Thousand yuan
	Total taxes of the same period last year grew (this year per month)	%
	Asset liability ratio	%
	Turnover of current assets	Time
	Cost profit margin	%

2.2.2 Data Preprocessing

(1) Missing data

Data were collected, the province of an index time series data missing points for the following three cases: missing the head, tail absent; halfway interrupts. To account for the missing data width of little time, generally only 1–8 months (with occasional exceptions lack of data of 15 months), under the assumption of stationary, we adopt the following additional interpolation.

① The head of the sequence is missing: The assumption that the value of the data x_n is missing, then the order can be $x = \cdots = x_{n-1} = x_n$;

② The tail sequence is missing: it is assumed that the value of the data x_k is missing, then the order can be $x_{k+1} = \cdots = x_N = x_k$;

③ The central part of the sequence is missing: If all the missing values x_p to x_q be missing, then the recursion formula is performed according to the following recursion formula [5]:

$$x_{p+1} = x_p + \frac{x_q - x_p}{q - p}, x_{p+2} = x_{p+1} + \frac{x_q - x_p}{q - p},$$

In a general way,

$$x_{p+i} = x_{p+i-1} + \frac{x_q - x_p}{q - p} \quad (1 \le i \le p - q - 1) \tag{1}$$

(2) Normalization of data

① The growth rate index: The different index types of statistical caliber, we need to be standardized. In this paper, we will draw on the method of data standardization in clustering analysis. The formula is as follows [6]:

$$x'_{ij}(t) = \frac{x_{ij}(t) - \bar{x}_{ij}}{s_{ij}} \tag{2}$$

② The main problem of this kind of index is that the size of the main problem is serious, and it is easy to cause the deviation. The problem can be solved by using maximum minimum value method [7]:

$$y'_{ij}(t) = \frac{y_{ij}(t) - \min\{y_{ij}(t)\}}{\max\{y_{ij}(t)\} - \min\{y_{ij}(t)\}} \tag{3}$$

The observation value of the index of the province at the moment is the standard result.

2.3 Model Establishment

In this part, two tasks are solved by establishing three models. After the establishment of the index model of electronic information industry development, according to 35 indicators are further calculated to measure the various provinces and autonomous regions of the information industry development index; then according to different time of various provinces and autonomous regions of the index changes, the establishment of time series forecasting model to predict the 2014 indices are calculated. Finally, aiming at the problem of the evaluation, using Bayesian classifier was evaluated, the Bayesian classifier can be different values of the provinces and cities in the annual classification, so as to reveal the electronic information industry development trend in various provinces and cities.

3 Text Clustering Algorithm Based on Factors Space

3.1 The Establishment of the Electronic Information Industry Development Index Model

For a province, considering the 35 indicators in a year's time length. The normalized difference between the various specific indicators X_k $(k = 1, 2, \ldots, 35)$ has a correlation, which can be calculated from any of the two indicators of covariance. Set two factors of the correlation sequence is:

$$\begin{pmatrix} X_i^{(1)} & X_i^{(2)} & \cdots & X_i^{(n)} \\ X_j^{(1)} & X_j^{(2)} & \cdots & X_j^{(n)} \end{pmatrix}$$

Its covariance is Cov_{ij}, and the covariance of an exponential form a covariance matrix:

$$C = \begin{pmatrix} Var(X_1) & Cov_{12} & \cdots & Cov_{1,35} \\ Cov_{12} & Var(X_2) & \cdots & Cov_{2,35} \\ \vdots & \vdots & \ddots & \vdots \\ Cov_{1,35} & Cov_{2,35} & \cdots & Var(X_{35}) \end{pmatrix}$$

Let $X_m = \sum_{i=1}^{35} X_i$, as C's i line: $Cov_{im} = \sum_{k=1}^{35} Cov_{ik}$; in the C, we have $Var_m = \sum_{i=1}^{6} Cov_{im}$.

Let $r_i = \frac{Cov_{im}}{Var_m}$, the ratio r_i of the volatility of a certain index to the total volatility can be used as the weight of the index. Naturally, the development of the electronic information industry in a province, the province's electronic information industry development index can be calculated as follows:

$$\alpha = \sum_{i=1}^{35} r_i X_i \tag{4}$$

3.2 Time Series Forecasting Model

Step 1 Fitting ARMA (2, 1) model

(A) Fitting AR (3) model
The model to be fitted is

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \alpha_t$$

Fitting with least square method φ_1, φ_2, φ_3, and in the AR(3) model, Inverse parameter $I_j = \varphi_j$, so we can calculate I_1, I_2, I_3;

(B) Estimated ARMA (2, 1) model parameters initial values:
As the ARMA(2, 1)model, we also knows that

$$I_j - \theta_1 I_{j-1} = 0, j > 2$$
$$I_3 - \theta_1 I_2 = 0 \Rightarrow \theta_1 = \frac{I_3}{I_2}$$

In this we can calculate that the φ_1, φ_2 in the ARMA(2, 1)model

$$\begin{cases} \varphi_1 = I_1 + \theta_1 \\ \varphi_2 = \theta_2 - \theta_1 I_1 + I_2 \end{cases}$$

(C) Using nonlinear least square method φ_1, φ_2, θ_1 Value and confidence interval, And sum of squares of residuals and RSS;

Step 2 Step 1 has been fitted out $ARMA(2n, 2n-1)$, $n = 2$, and then to make the $n = n + 1$ fitting $ARMA(2n, 2n-1)$ model;

Step 3 Fitting model with F statistics. If passed, continue Step 1; if not through, into Step 4.
The statistical model of the hypothesis test is as follows [8]:

$$H_0: ARMA(n_2, m_2) \text{ is much more less than } ARMA(n_1, m_1) \text{in } RSS\ vs$$
$$H_1: ARMA(n_2, m_2) \text{ is not much less than } ARMA(n_1, m_1) \text{in } RSS$$

let $A_0 = RSS[ARMA(n_1, m_1)]$, $A_1 = RSS[ARMA(n_2, m_2)]$, Then the test statistic is:

$$F = \frac{N - (n_2 + m_2)}{(n_2 + m_2) - (n_1 + m_1)} \frac{A_1 - A_0}{A_0} \sim F((n_2 + m_2) - (n_1 + m_1), N - (n_2 + m_2))$$

N is the total data volume, the statistical model of the rejection region is:

$$W = \{F \leq F_{0.95}((n_2 + m_2) - (n_1 + m_1), N - (n_2 + m_2))\}$$

Step 4 Check whether the confidence interval is 0, if not, then apply $ARMA(2n, 2n-1)$;

Step 5 Use F to prove $ARMA(2n, 2n-1)$ and $ARMA(2n-1, 2n-2)$, in this time [9]:

$$F = \frac{N-(4n-1)}{(4n-1)-(4n-3)} \frac{RSS[ARMA(2n-1,2n-2)] - RSS[ARMA(2n,2n-1)}{RSS[ARMA(2n,2n-1)}$$

Rejection region is $W = \{F \le F_{0.95}(2, N-(4n-1))\}$
If $F \in W$, into Step 6;
If F $F \notin W$, into Step 7.

Step 6 Fitting $m < 2n-2$ model $ARMA(2n-1, m)$, use F to prove, do it until the model have the less Parameter;

Step 7 Fitting $m < 2n-1$ model $ARMA(2n, m)$, use F to prove, do it until the model have the less Parameter:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \alpha_t \tag{5}$$

3.3 Bayesian Classifier Model

Set up a class collection, classification, electronic information industry related indicators, there is a. Determine the mapping rule, and make an arbitrary, and only one such that is established. (Without considering the fuzzy mathematics in fuzzy set) which is called set of categories, where each element is a category, called item collection, of which every element is a to be classified items, known as classifier. The task of classification is to construct a classifier.

According to Bayes formula, the posterior probability density function [10]

$$P(\omega_i|x) = \frac{P(x|\omega_i)P(\omega_i)}{P(x)} \tag{6}$$

The prior probability is calculated by the training data directly. Then by the full probability formula,

$$P(x) = \sum_i P(x|\omega_i)P(\omega_i) \tag{7}$$

let $\kappa_0 = 1/P(x)$, as a constant, the posterior probability density function is transformed to

$$P(\omega_i|x) = \kappa_0 P(x|\omega_i)P(\omega_i) \tag{8}$$

From the formula (8), it can be known that the electronic information industry in a certain province is a probability of a certain level, and it is very important to be determined by the probability of a certain kind of province. According to the classification method of Naive Bayesian, the electronic information industry index of the province to be inquired is the category of the index, which is satisfied.

$$P(\omega_k|x^*) = \max\{P(\omega_i|x^*)\} \quad (9)$$

For the class conditional probability density function, we make the following discussion:

(1) The estimation of normal distribution
In fact, the construction of conditional probability density function must be based on a variety of methods, the purpose is to make the actual situation close to the internal law. Normal distribution discrimination is reasonable in mathematics, widely and only the mean and variance of the two parameters, practical significance, unimodality of electronic information industry more obvious, normal estimation is very clear and reasonable. The probability density function of the class can be given by the multivariate normal distribution:

$$P(x|\omega_i) = \frac{1}{(2\pi)^{n/2}|Cv_i|^{1/2}} e^{[-\frac{1}{2}(x-\mu)^T Cv_i^{-1}(x-\mu)]} \quad (10)$$

In there, $x = [x_1, x_2, \ldots x_n]^T$, is an N-dimensional feature vector;
$\mu = [\mu_1, \mu_2, \ldots \mu_n]^T$, for the n-dimensional mean vector;
sample data at all levels within in is $\mu_i = E[x_i]$.
Cv is Covariance matrix for a certain class, $Cv = E[(x-\mu)(x-\mu)^T]$.

(2) The kernel density smoothing estimation
Kernel density estimation is a relatively perfect statistical field. The objective of this technique is to describe the data distribution with function [2]. Because of a small sample in a class, we use the following kernel density smoothing estimation as an auxiliary method of calculation.
For a given random sample $x = (x_1, x_2, \ldots x_n)$, the kernel density estimation of the probability density function is estimated.

$$P(x, h|\omega_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}) \quad (11)$$

The kernel density estimation function is the bandwidth, and the kernel density estimation is close to the true density value. The specific ideas can be seen in the one-dimensional showing (Fig. 2).
Bandwidth is a scaling factor, and the probability distribution of the control is near to the sample, and the Matlab toolbox is based on the appropriate values.

(3) The calculation of training sample index matrix *Property*:
Bayesian method in solving high dimension and does not fall into high dimensional crisis, in high virial more relevant indicators of the impact will make inspection probability density function is more accurate, when using the naive Bayesian classifier, following initial treatment of the electronic information industry in various indicators: (1) the data in 2003.2-2012.12 average

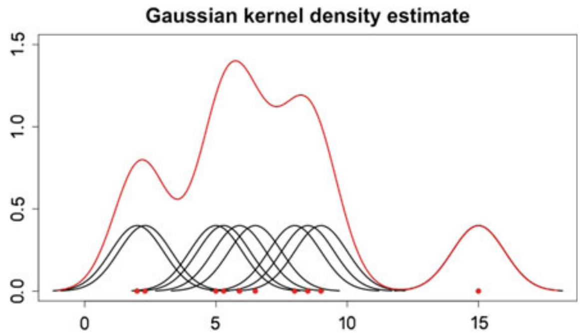


Fig. 2 Gaussian kernel density estimate

as the index attribute value; (2) neglect not statistical indicators of the impact. Training matrix *Property* [11].

Using Matlab toolbox function *Naive Bayes*:

$$ObjBayes = NaiveBayes.fit(meas, Class);$$
$$pre\,0 = ObjBayes.predict(meas);$$

The distribution of the above class conditional function is not defined as normal distribution in Matlab. After the training, the value of pre0 and species was obtained, and the number of the false product was obtained.

Kernel density smoothing is used to call distribution.

4 Model Solution

4.1 *Calculation and Forecast of the Development α Index of the Electronic Information Industry in the Provinces*

The predicted time series model uses MA (2) model, and the results are as follows (Table 2).

4.2 *Bayesian Hierarchical System*

The development trend of the electronic information industry, especially the medium level of development, is particularly important in the macro planning. We want to have a sound classification system for the electronic information industry [12], so that in the decision process to play a role in guiding the

Table 2 α index of the provinces

Province	2012	2013	2014 (forecast)
Shanghai	1.06	1.1	1.128316004
Beijing	1.04	1.1	1.155146063
Zhejiang	0.99	1.07	1.032690455
Tianjin	0.97	1.08	1.053904092
Jiangsu	0.97	1.03	1.067189905
Fujian	0.965	1.02	1.062977618
Guangdong	0.96	1.04	1.047656531
Shandong	0.88	0.89	0.957184512
Hubei	0.86	0.87	0.952486767
Sanxi	0.83	0.84	0.861571606
Liaoning	0.82	0.85	0.912531071
Hainan	0.815	0.84	0.881510058
NingXia	0.81	0.83	0.911134186
Shanxi	0.805	0.825	0.86980764
Hebei	0.80	0.82	0.83437254
Guangxi	0.77	0.80	0.79591836
Jilin	0.765	0.795	0.877113223
Sichuan	0.76	0.79	0.85142059
Anhui	0.73	0.76	0.799123576
Xinjiang	0.71	0.74	0.813524528
Jiangxi	0.705	0.73	0.788789715
Chongqing	0.70	0.72	0.729843176
Inner Mongolia	0.64	0.66	0.718758299
Yunnan	0.635	0.68	0.752815098
Heilongjiang	0.63	0.63	0.63
Hunan	0.625	0.64	0.72816767
Gansu	0.62	0.61	0.623653209
Henan	0.61	0.65	0.634329908
Qinghai	0.605	0.61	0.701720254
Guizhou	0.55	0.55	0.55

development of the industry [13]. According to the idea of Naive Bayesian classifier, the following Bayesian classifier is built (Tables 3 and 4).

According to the “report” in the province of information index size, we will be the province’s electronic information industry development level is divided into 5 levels, as follows:

Through the Bayesian classifier, we can make a comparison of the development level of the electronic information industry in 2013 and 2014.

Table 3 Levels of electronic information industry development

Development level of electronic information industry	Information index in the "Report"	Representative region	Range of α index
To develop	≥100	Shanghai, Beijing	≥1
More developed	90–100	Zhejiang, Tianjin, Jiangsu, Fujian, Guangdomg	0.9–1
Development	70–80	Shandong, Hubei, Shanxi, Liaoning, Hainan, Ningxia, Shanxi	0.8–0.9
To be improved	60–70	Hebei, Guangxi, Jilin, Sichuan, Anhui, Xinjiang, Jaingxi, Chongqing, Neimenggu, Yunnan, Heilongjiang, Hunan, Gansu, Henan, Qinghai	0.6–0.8
Less developed	Less than 60	Guizhou	<0.6

Table 4 Comparison of development level of electronic information industry (2012–2014)

Development level of electronic information industry	2012	2013	2014 (forecast)
To develop	Shanghai, Beijing	Shanghai, Beijing, Zhejiang, Tianjin, Jiangsu, Fujian, Guangdong	Shanghai, Beijing, Zhejiang, Tianjin, Jiangsu, Fujian, Guangdong
More developed	Zhejiang, Tianjin, Jiangsu, Fujian, Guangdong	–	Shandong, Hubei, Liaoning, Ningxia, Hianan
Development	Shandong, Hubei, Shanxi, Liaoning, Hainan, Ningxia, Shaanxi	Shandong, Hubei, Shaanxi, Liaoning, Hainan, Ningxia, Shanxi, Hebei, Guangxi	Shanxi, Hebei, Xinjiang, Jilin, Sichuan, Shannxi
To be improved	Hebei, Guangxi, Jilin, Sichuan, Anhui, Xinjiang, Jiangxi, Chongqing, Neimenggu, Yunnan, Heilongjiang, Hunan, Gansu, Henan, Qinghai	Jilin, Sichjuan, Anhui, Xinjiang, Jiangxi, Chongqing, Neimenggu, Yunnan, Heilongjiang, Hunan, Gansu, Henan, Qinghai	Anhui, Jiangxi, Chongqing, Neimenggu, Yunnan, Heilongjiang, Hunan, Gansu, Henan, Qinghai
Less developed	Guizhou	Guizhou	Guizhou

5 Conclusion

According to the ranking situation, the average value of the index is used to generate the Bayesian classification system. Enter the time series of the forecast value into the system, a new ranking will be presented, expressed as category.

(a) According to before and after ranking, industry trends can be compared.
(b) For a certain province, according to the fluctuation of the province class, it can be analyzed that the index is sensitive.
(c) For other provinces or regions of the industry, the input of the Bayesian classification system, the gap can be compared.

Acknowledgements This research was supported by Information and Computing Science Outstanding Talent Training Project of Guangdong Province (No.20153324).

References

1. Zhong, Y., Li, Y., Duan, Y.: Application research on greed strategy-based GA in OCS model in universities. Adv. Intell. Soft Comput. **147**, 379–389. ISSN 1867–5662 (Springer) (2012)
2. Zhong, Y., Xiang, Y.: Research and application of parallel fuzzy dynamic time warping algorithm in music retrieval. Adv. Intell. Soft Comput. **147**, 505–513. ISSN 1867–5662 (Springer) (2012)
3. Zhong, Y., Xiang, Y.: Design and realization of music recognition based on speech recognition. IEEE Netw. Comput. Inf. Security **2** (2011)
4. Zhong, Y., Li, H.: The upgrade of topological group based on a new hyper-topology. Fuzzy Inf. Eng. **3**(2), 213–224. ISSN 1616–8658 (Springer) (2011)
5. Zhong, Y., Zheng, J., Cao, B.: The structure and expression of power group. Fuzzy Inf. Eng. **2** (2). ISSN 1616–8658 (Springer) (2010)
6. Zhong, Y., Li, B.: Design and realization of FCE optimized model in DSS. IEEE. Comput. **1** (2009)
7. Zhong, Y., Mi, H., Zhong, W., Zhu, J.: Simulation and optimization of the flow in ERP system for MMI. IEEE. Comput. **4**, 2182–2186 (2009)
8. Zhong, Y., Deng, H., Chen, H.: A fuzzy logic based mobile intelligent system for effectively evaluating and selecting hotels in tourism. IEEE. Web Inf. Syst. Min. 733–737 (2009)
9. Zhong, Y., Mi, H., Chen, H.: Design and optimization algorithm of VCS under security protection. IEEE. Web Inf. Syst. Min. 552–556 (2009)
10. Zhong, Y., Lin, Z., Chen, R.: Realization and optimization of the flow in ERP system for RCI. IEEE. Web Inf. Syst. Min. 792–796 (2009)
11. Zhong, Y.: The design of a controller in fuzzy petri net. Fuzzy Optim. Decis. Making **7**(4), 399–408. ISSN 1568–4539 (Springer) (2008)
12. Zhong, Y.: The optimization for location for large commodity's regional distribution center. Adv. Soft Comput. **40**, 969–979. ISSN 1615-3871 (Springer) (2007)
13. Yang, J., Zhao, Y., Ding, W.X.: Imputation methods of missing data in sampling survey. Appl. Stat. Manag. 2008 (05)

Fuzzy Formulation of the Lee-Carter Model for the Mortality Forecasting with Age-Specific Enhancement

Yefu Kou

Abstract In this paper, we consider the fuzzy method of Koissi and Shapiro [12] and apply the Lee-Carter model with age-specific enhancement of Renshaw and Haberman [13–15]. The proposed fuzzy formulation of the extended Lee-Carter model is exercised based on mortality data like the mortality of China between 1994 and 2008. We also obtain the predictive mortality rate. The comparative advantages of our proposed fuzzy formulation of the extended Lee-Carter model, relative to the classical Lee-Carter model, are analyzed and discussed.

Keywords Lee-Carter model · Fuzzy modeling · Mortality forecasting

1 Introduction

Apart from being a country with the largest population in the world, the population of China is experiencing a rapid aging over the past half a century. According to the China Population and Employment Statistics yearbook (2009), the proportion of population aged 65 or older was only 4.4 % in 1953 and 4.91 % in 1982, but increased to almost twice (8.3 %) in 2008, and by 2030 it will be more than double to 22 %. With the rapid growth in aging and the trends in improving mortality among the elderly, such phenomenon exerts significant challenges to the public pension plans as well as private pension funds and life insurers, particularly in China. Consequently, there is a demand in a stochastic mortality model which adequately projects the mortality/longevity trends of the China population.

In this paper there are two main contributions. The first contribution is to consider a fuzzy formulation of the modified Lee-Carter [11] model analyzed in Renshaw and Haberman [16–18], thus extending the work of Koissi and Shapiro [10] which only consider the fuzzy formulation of the basic Lee-Carter model.

Y. Kou (✉)
China Institute for Actuarial Science, Central University of Finance and Economics, Beijing 100081, People's Republic of China
e-mail: kouyefu@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_18

As argued in Koissi and Shapiro [10] that the fuzzy formulation of the Lee-Carter model has the advantage of addressing violations of the Lee-Carter assumption of constant error variance across age. The vagueness in the definition of the error term in the Lee-Carter model further supports the use of a fuzzy implementation.

In addition to mortality modeling, the fuzzy set theory has also found its usefulness in a variety of other insurance applications, particularly in calculating insurance claim reserving [1, 5], insurance underwriting [7, 12], classification of insurance risk [6, 9], pricing property-liability insurance [4, 19], adjusting insurance rate [22]. See also Shapiro [20, 21] and the monograph by Ostaszewski [13] on applications of fuzzy logic to actuarial science.

The second contribution of the paper is to implement the fuzzy formulation of the basic Lee-Carter model and the extended Lee-Carter model (with age-specific enhancement) on the China data. The comparative advantages of our proposed fuzzy formulation of the extended Lee-Carter model, relative to the classical Lee-Carter model, are analyzed and discussed. This is of interest as most of the published results related to the Lee-Carter model have been applied on the developed countries, instead of the developing country such as China.

The remaining of the paper is organized as follows. Section 2 reviews the Lee-Carter model, a variant of the Lee-Carter considered by Renshaw and Haberman [16–18] and Booth et al. [3] and fuzzy set theory. Section 3 explains the fuzzy formulation of the modified Lee-Carter model. Section 4 implements the fuzzy formulation of the classical Lee-Carter model and the modified Lee-Carter model on China population data and some concludes of the paper.

2 Review of the Lee-Carter Model and Fuzzy Set Theory

We begin by recalling the basic Lee-Carter (LC) model, which is given by

$$\ln m_{xt} = a_x + b_x k_t + e_{xt}, \quad x = x_1, \ldots, x_N \; and \; t = t_1, t_1 + 1, \ldots, t_1 + T - 1 \tag{1}$$

where m_{xt} represents the matrix of the central death rates at age x in year t, N is the number of age-groupings, and T is the number of years. a_x, b_x, and k_t are age and time dependent parameters. The error term ε_{xt} is Gaussian $\varepsilon_{xt} \sim N(0, 1)$ and (1) is solved under the constraints that $\sum_t k_t = 0; \quad \sum_x b_x = 1$.

An extension of the LC model is possible by including higher order terms, thereby allowing for greater flexibility in forecasting changes. In this case, the modified LC model has the following form:

$$\log m_{xt} = a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)} + \cdots + b_x^{(r)} k_t^{(r)} + \varepsilon_{xt}, \tag{2}$$

In this model, the age-time interaction term $b_x^{(i)} k_t^{(i)}$ is referred to as the i-th term of the rank r approximation (see Booth et al. [2]). Renshaw and Haberman [22–24] investigate the above modified Lee-Carter model with age-specific enhancement for mortality forecasts by considering r = 2. Note that when r = 1, the above model reduces to the basic LC model. In this paper, we are interested in the fuzzy formulation of the modified LC model (2) with r = 2. We denote the resulting model as LC2.

We now provide some notations and concepts related to fuzzy set theory. The description below is drawn largely from Koissi and Shapiro [11].

A crisp set C, is given by a characteristic function $\chi_c(x)$ such that $\chi_c(x) = 1$ if x is in C; otherwise, $\chi_c(x) = 0$. Fuzzy sets, on the other hand, generalize non-fuzzy or crisp sets. The objects in fuzzy sets can belong to a given set to different degrees and in this sense, they generalize the notion of set to allow partial membership in a set:

Definition 1 A fuzzy subset A (over a reference set X) is a function μ_A on X that takes values in the unit-interval [0, 1]: $\mu_A \rightarrow [0, 1]$. The μ_A is called the membership function of A and for any x in X, $\mu_A(x)$ in [0,1] denotes the grade of membership of x in A.

Fuzzy sets can also be characterized by their λ-cuts, which is defined as

$$A_\lambda = \{x \in X : \mu_A(x) \geq \lambda\},\ 0 \leq \lambda \leq 1.$$

Here λ can be interpreted as the confidence level.

When the reference set is the real line *R*, then the set A is called fuzzy number. Fuzzy numbers are characterized by their membership functions, which can be triangular, trapezoidal, Gaussian, etc. Here we consider the triangular membership function.

Definition 2 [32] Let $\tilde{A} = (a, l_a, r_a)$ be a triangular fuzzy number with center $a \in R$ and left and right spreads (l_a, r_a). Then its characteristic can be denoted by a membership function $\mu_{\tilde{A}}(x)$ or α-cut set:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x - a + l_a}{l_a}, & a - l_a < x \leq a \\ \frac{a + r_a - x}{r_a}, & a < x \leq a + r_a \\ 0, & otherwise \end{cases}$$
$$A_\alpha = [\underline{A}(\alpha), \bar{A}(\alpha)] = [a - l_a(1 - \alpha), a + r_a(1 - \alpha)].$$

Figure 1 provides a graphical representation of a triangular fuzzy number. As an example, if you believe that the inflation rate of the loss cost in the next two years is about 2 %, its deviation is not more than 1 %. Then this is an example of a triangular fuzzy number which can be expressed as (0.02, 0.01, 0.01). In the special case where $l_a = r_a$, the triangular fuzzy number becomes a symmetric triangular fuzzy number. In this case, the symmetric triangular fuzzy number is denoted by $A = (a, l)$.

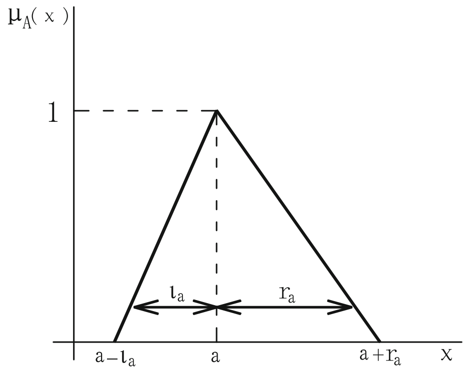


Fig. 1 Triangular fuzzy number: $\tilde{A} = (a, l_a, r_a)$

3 Fuzzy Formulation of the LC Model with Age-Specific Enhancement

A fuzzy formulation of the Lee-Carter model is

$$\tilde{Y}_{x,t} = \tilde{A}_x \oplus \sum_{i=1}^{2} \left(\tilde{B}_x^{(i)} \otimes \tilde{K}_t^{(i)} \right), \; x = x_1, \ldots, x_N \quad t = t_1, t_1 + 1, \ldots, t_1 + T - 1 \quad (3)$$

where $\tilde{Y}_{x,t}$ are known fuzzy log-central death rates and $\tilde{A}_x$, $\tilde{B}_x^{(i)}, \tilde{K}_t^{(i)}$ are unknowns. The parameter $\tilde{A}_x$ is the fuzzy representation of the average age-specific pattern of mortality, $\tilde{K}_t^{(i)}$ represents the fuzzy formulation of the general mortality level and $\tilde{B}_x^{(i)}$ captures the decline in mortality at age x. In the above fuzzy formulation, $\oplus$ and $\otimes$ are fuzzy operators and by assuming $\tilde{A}_x$, $\tilde{B}_x^{(i)}, \tilde{K}_t^{(i)}$ are symmetric triangular fuzzy numbers such that:

$$\tilde{A}_x = (a_x, \alpha_x), \, \tilde{B}_x^{(i)} = \left(b_x^{(i)}, \beta_x^{(i)} \right), \, \tilde{K}_t^{(i)} = \left(k_t^{(i)}, \delta_t^{(i)} \right) \quad (4)$$

Then the fuzzy "addition" and "multiplication" reduce to

$$\tilde{A}_x \oplus \tilde{B}_x^{(i)} = \left(a_x + b_x^{(i)}, \max\left(\alpha_x, \beta_x^{(i)} \right) \right)$$
$$\tilde{A}_x \otimes \tilde{B}_x^{(i)} = \left(a_x b_x^{(i)}, \max\left(\alpha_x \left| b_x^{(i)} \right|, \beta_x^{(i)} |a_x| \right) \right).$$

Consequently (3) becomes

$$\begin{aligned}\tilde{Y}_{x,t} &= \tilde{A}_x \oplus \sum_{i=1}^{2}\left(\tilde{B}_x^{(i)} \otimes \tilde{K}_t^{(i)}\right)\\ &= \left(a_x + b_x^{(1)}k_t^{(1)}, \max\left(\alpha_x, \left|b_x^{(1)}\right|\delta_t^{(1)}, \beta_x^{(1)}\left|k_t^{(1)}\right|\right)\right) \oplus (\tilde{B}_x^{(2)} \otimes \tilde{K}_t^{(2)})\\ &= \left(a_x + b_x^{(1)}k_t^{(1)} + b_x^{(2)}k_t^{(2)}, \max\left(\alpha_x, \left|b_x^{(1)}\right|\delta_t^{(1)}, \beta_x^{(1)}\left|k_t^{(1)}\right|, \left|b_x^{(2)}\right|\delta_t^{(2)}, \beta_x^{(2)}\left|k_t^{(2)}\right|\right)\right).\end{aligned} \tag{5}$$

This is the formulation of Fuzzy-LC with age-specific enhancement. It should be emphasized that in the above formulation, the log-central death rate for age-group x in year t is a symmetric triangular fuzzy number, instead of exactly $\ln(m_{x,t})$ as in the traditional non-fuzzy formulation. Solving (5) involves two key steps. The first step is to fuzzily $Y_{x,t} = \ln(m_{x,t})$. The second and final step is to estimate the fuzzy-LC parameters.

Step 1: Fuzzification of $Y_{x,t} = \ln(m_{x,t})$

The objective of this step is to fuzzily the log-central death rates $Y_{x,t}$. For simplicity we assume that the fuzziness is captured by the symmetric triangular membership function. By introducing

$$\tilde{Y}_{x,t} = (y_{x,t}, e_{x,t}), \quad \tilde{A}_1 = (c_{0x}, s_{0x}), \quad \tilde{A}_1 = (c_{1x}, s_{1x}), \quad \text{and} \quad \tilde{A}_2 = (c_{2x}, s_{2x}),$$

We have

$$(y_{x,t}, e_{x,t}) = (c_{0x}, s_{0x}) + (c_{1x}, s_{1x}) \times t^{(1)} + (c_{2x}, s_{2x}) \times t^{(2)}. \tag{6}$$

The above formulation can further be simplified by noting $t^{(1)} = t^{(2)}$ so that

$$\begin{aligned}(y_{x,t}, e_{x,t}) &= (c_{0x}, s_{0x}) + [(c_{1x}, s_{1x}) + (c_{2x}, s_{2x})] \times t\\ &= (c_{0x}, s_{0x}) + (c_{3x}, s_{3x}) \times t\end{aligned} \tag{7}$$

where $(c_{3x}, s_{3x}) = (c_{1x}, s_{1x}) + (c_{2x}, s_{2x})$.

Consequently, the remaining task is to find centers c_{0x}, c_{3x} and spreads s_{0x}, s_{3x}. As argued in Koissi and Shapiro [10], the centers c_{0x}, c_{3x} are easily found by fitting the ordinary least-squares regression to $Y_{x,t} = c_{0x} + c_{3x} \times t$ where $Y_{x,t}$ are the observed log-central deaths. The spreads s_{0x}, s_{3x} are then obtained based on minimum fuzziness criterion which consists of simultaneously minimizing the spreads of the estimated

$$\hat{\tilde{Y}}_{x,t} = (\hat{c}_{0x}, \hat{s}_{0x}) + (\hat{c}_{3x}, \hat{s}_{3x}) \times t \tag{8}$$

and requiring each death rate $\tilde{Y}_{x,t}$ to fall within the estimated $\hat{\tilde{Y}}_{x,t}$ at a level h. The latter requirement is can fulfilled by ensuring that $\mu(\tilde{Y}_{x,t} \subseteq \hat{\tilde{Y}}_{x,t}) \geq h,\ h \in [0, 1]$. The level h, which measures the degree of fit of the estimated model to the given data, is a user input.

Step 2: Estimating the parameters of fuzzy-LC with age-specific enhancement

The objective of this step is to determine $\tilde{A}_x$, $\tilde{B}_x^{(i)}, \tilde{K}_t^{(i)}$, $i = 1, 2$. We achieve this task by minimizing the square of the distance between $\tilde{A}_x \oplus \sum_{i=1}^{2} (\tilde{B}_x^{(i)} \otimes \tilde{K}_t^{(i)})$ and $\tilde{Y}_{x,t}$. Here we adopt Diamond [9, p. 143] distance measure, which for two fuzzy sets $\tilde{A}_1 = (a_1, \alpha_1)$ and $\tilde{A}_2 = (a_2, \alpha_2)$, the distance is measured by

$$D_{LR}(\tilde{A}_1, \tilde{A}_2)^2 = (a_1 - \alpha_2)^2 + [(a_1 - \alpha_1) - (a_2 - \alpha_2)]^2 + [(a_1 + \alpha_1) - (a_2 + \alpha_2)]^2. \tag{9}$$

In summary, the estimation of the fuzzy-LC parameters boils down to solving the following minimization problem:

$$\text{Minimize} \sum_x \sum_t D_{LR}[\tilde{A}_1 \oplus \sum_{i=1}^{2} (\tilde{B}_x^{(i)} \otimes \tilde{K}_t^{(i)}), \tilde{Y}_{x,t}]^2 \tag{10}$$

where

$$\begin{aligned} F &= D_{LR}[\tilde{A}_x \oplus \sum_{i=1}^{2} (\tilde{B}_x^{(i)} \otimes \tilde{K}_t^{(i)}), \tilde{Y}_{x,t}]^2 \\ &= D_{LR}[\left(a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)}, \max\left(\alpha_x, \left|b_x^{(1)}\right| \delta_t^{(1)}, \beta_x^{(1)} \left|k_t^{(1)}\right|, \left|b_x^{(2)}\right| \delta_t^{(2)}, \beta_x^{(2)} \left|k_t^{(2)}\right|\right)\right), (y_{x,t}, e_{x,t})]^2 \\ &= (a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)} - y_{x,t})^2 \\ &\quad + [\left(a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)} - \max\left(\alpha_x, \left|b_x^{(1)}\right| \delta_t^{(1)}, \beta_x^{(1)} \left|k_t^{(1)}\right|, \left|b_x^{(2)}\right| \delta_t^{(2)}, \beta_x^{(2)} \left|k_t^{(2)}\right|\right)\right) - (y_{x,t} - e_{x,t})]^2 \\ &\quad + [\left(a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)} + \max\left(\alpha_x, \left|b_x^{(1)}\right| \delta_t^{(1)}, \beta_x^{(1)} \left|k_t^{(1)}\right|, \left|b_x^{(2)}\right| \delta_t^{(2)}, \beta_x^{(2)} \left|k_t^{(2)}\right|\right)\right) - (y_{x,t} + e_{x,t})]^2 \end{aligned} \tag{11}$$

$$\text{subject to} : \begin{cases} a_x = (1/T) \sum_t y_{x,t} \\ \sum_x b_x^{(i)} = 1, \sum_{t=t_1}^{t_n} k_t^{(i)} = 0 \quad i = 1, 2 \end{cases}. \tag{12}$$

MATLAB can be used to obtain the optimal parameters α_x, $b_x^{(1)}$, $\beta_x^{(1)}$, $k_t^{(1)}$, $\delta_t^{(1)}$, $b_x^{(2)}$, $\beta_x^{(2)}$, $k_t^{(2)}$ and $\delta_t^{(2)}$.

Once $b_x^{(1)}$, $k_t^{(1)}$, $b_x^{(2)}$, and $k_t^{(2)}$ are estimated, $k_t^{(2)}$ is further adjusted to $k_t^{(2)'}$ so that the actual total deaths and the total expected deaths for each t matches. In other words, the parameter estimates satisfy

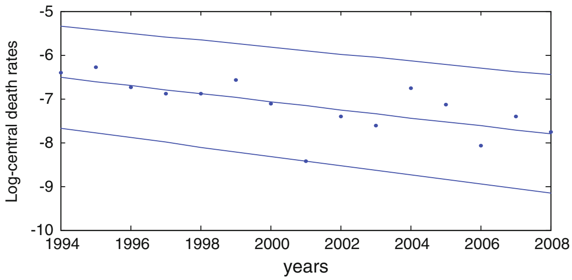


Fig. 2 Fuzzy least-squares regression of log-central death rates (Male, X = 10, age group 27–29) with age-specific enhancement

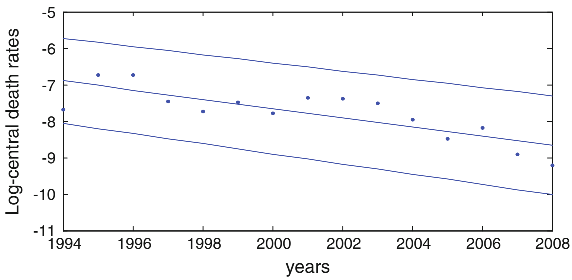


Fig. 3 Fuzzy least-squares regression of log-central death rates (Female, X = 10, age group 27–29) with age-specific enhancement

$$\sum_{x=x_1}^{x_k} d_{xt} = \sum_{x=x_1}^{x_k} n_{xt} \exp(\hat{a}_x + \sum_{i=1}^{2} \hat{b}_x^i \hat{k}_t^i), \forall t \tag{13}$$

where d_{xt} is the number of death of people at age x and time t, n_{xt} is the total number of people at age x and time t.

Then we forecast $k_t^{(1)}$, $k_t^{(2)}$ using ARIMA and get:

$$k_t^{(1)}, k_t^{(2)} \quad t = t_0 + T, t_0 + T + 1, \ldots, t_0 + T + s$$

Lastly, forecast the mortality through:

$$\hat{y}_{xt} = a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)}, \quad t = t_0 + T, t_0 + T + 1, \ldots, t_0 + T + s$$

We can get Male and Female c_{0x}, c_{1x}, s_{0x}, s_{1x} with age-specific enhancement (Figs. 2 and 3).

4 Mortality Forecasting: China Data

This paper uses 15 yearly observations of age-specific death rates for both males and females in China from 1994 to 2008, covering ages 0 to 89. These data are provided by the China Population Statistical Yearbooks and the China Statistical

Yearbooks compiled by the National Bureau of Statistics of China. We implement the fuzzy formulation of Lee-Carter with r = 1 and r = 2 by partitioning the age into 30 groups consist of [0,2], [3,5], [6,8], …, [87,89]. In other words, we have t = 1, 2, …, 15 and x = 1, 2, … 30.

Estimating fuzzy-LC parameters with China death rates data.

To estimate fuzzy-LC parameters, we need to define the formulation as follows, because of the min function in (10).

Assume:

$$r(x,t,i) = \begin{cases} \alpha_x & i = 1 \\ |b_x|\delta_t & i = 2 \\ \beta_x|k_t| & i = 3 \end{cases} \tag{14}$$

Then (13) become to the formulation of function g and function f:

$$g(r) = 2r^2 - 4e_{x,t}r \tag{15}$$

$$f(b_x, k_t) = 3(b_x k_t)^2 + 6a_x b_x k_t - 6b_x k_t y_{xt} + (3a_x^2 + 3y_{xt}^2 - 6a_x y_{xt} + 2e_{x,t}^2) \tag{16}$$

Then (10) changes to: Minimize $\sum_x \sum_t [f(b_x, k_t) + g(\max(r))]$.

Then we estimate the parameters using MATLAB.

The calculation methods of a_x, b_x, k_t are same as the fuzzy formulation of Lee-Carter with r = 1. Then we get male a_x, b_x, k_t and female a_x, b_x, k_t with China death rates data (Tables 1 and 2).

Then we forecast k_t from 2009 to 2050 with ARIMA. We select ARIMA (0,2,1) for male and ARIMA (1,2,0) for female using ADF test and R test. The results are given in Figs. 4 and 5.

Lastly, forecast the mortality from 2009 to 2050 with:

$$\hat{y}_{xt} = a_x + b_x k_t, \quad t = t_0 + T, t_0 + T + 1, \ldots, t_0 + T + s$$

Based on the formula of spread:

$$e_{x,t} = s_{0x} + s_{1x} \times t, \mathrm{t} = 1, 2, 3, \ldots 57.$$

We can calculate the spread of mortality. The calculation methods and results of the spread are same as the fuzzy formulation of Lee-Carter with r = 1 (Table 3).

Then we can draw the mortality forecasting figure (Figs. 6 and 7):

Table 1 Male and female a_x, b_x with China death rates data

x	Male		Female		x	Male		Female		X	Male		Female	
	a_x	b_x	a_x	b_x		a_x	b_x	a_x	b_x		a_x	b_x	a_x	b_x
1	−5.9045	0.1008	−5.6301	0.0499	11	−6.9839	0.0366	−7.7915	0.0357	21	−4.4171	0.0334	−4.9927	0.0256
2	−7.8870	0.0687	−7.9487	0.0273	12	−6.6554	0.0250	−7.5381	0.0229	22	−4.1013	0.0233	−4.6117	0.0257
3	−8.0049	0.0357	−8.1851	0.0259	13	−6.6136	0.0388	−7.3903	0.0177	23	−3.9145	0.0226	−4.3010	0.0276
4	−8.0594	0.0423	−8.5195	0.0425	14	−6.3345	0.0115	−6.9859	0.0366	24	−3.4409	0.0235	−3.9672	0.0254
5	−7.9919	0.0105	−8.4866	0.0226	15	−6.1151	0.0144	−6.9676	0.0295	25	−3.2706	0.0269	−3.7037	0.0266
6	−8.0640	0.0330	−8.4653	0.0434	16	−5.8784	0.0171	−6.7058	0.0242	26	−2.9688	0.0254	−3.3271	0.0229
7	−7.8281	0.0871	−8.0454	0.0734	17	−5.5493	0.0292	−6.1582	0.0419	27	−2.6537	0.0214	−3.0677	0.0228
8	−7.4323	0.0477	−8.0518	0.0653	18	−5.2739	0.0181	−5.9483	0.0205	28	−2.4046	0.0337	−2.6774	0.0237
9	−7.3403	0.0374	−7.9656	0.0747	19	−5.1053	0.0246	−5.7489	0.0214	29	−2.0912	0.0285	−2.4236	0.0181
10	−7.1610	0.0421	−7.7790	0.0560	20	−4.7481	0.0239	−5.2503	0.0287	30	−1.7503	0.0168	−2.1411	0.0214

Table 2 k_t with China death rates data

T		1	2	3	4	5	6	7	8	9	10	11	12	13	14
k_t	Male	11.52	17.37	10.18	8.39	5.98	5.85	3.87	0.63	−0.89	−2.50	−4.00	−4.46	−10.57	−15.68
	Female	13.93	21.64	10.26	8.18	5.66	4.76	5.16	1.41	−2.32	−0.07	−4.72	−4.57	−16.36	−12.99

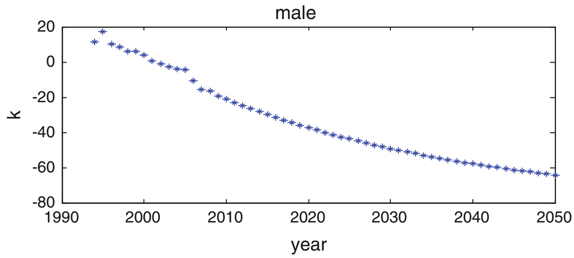


Fig. 4 Male k_t with China death rates data

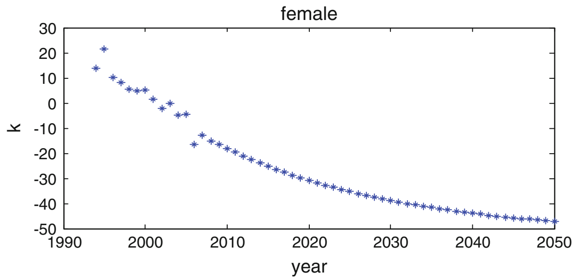


Fig. 5 Female k_t with China death rates data

In Figs. 6 and 7, the middle curve is the forecasting mortality; the upper and lower curves are the upper limit and lower limit of forecasting mortality.

Then we can calculate the life expectancy and compare them with the original data (Figs. 8 and 9).

We summarize the results of life expectancy:

In 2008, the expectancy of male is 79.07, the expectancy of female is 81.56.

In 2050, the expectancy of male is 85.89, the expectancy of female is 85.94.

The expectancy of male increased 6.82, the expectancy of female increased 4.08, and the difference of male and female decreased to 0.05 from 2.79.

Because the fuzzificating of $Y_{x,t}$ is same to single factor fuzzy-LC, so we can use the results of single factor fuzzy-LC. So, We estimating parameters of the fuzzy-LC with age-specific enhancement directly using MATLAB (Table 4):

Keep $k_t^{(1)}$ stable, justify $k_t^{(2)}$ based on (13), get $k_t^{(2)'}$ (Table 5):

Then we forecast the $k_t^{(1)}$ and $k_t^{(2)'}$ to 2050 with ARIMA.

(1) $k_t^{(1)}$: We select ARIMA (0,2,1) for male and ARIMA (0,1,1) for female using ADF test and R test.

Table 3 Male and female (x = 10) mortality forecasting

Year	$\ln \tilde{Y}_{x,t}$							
	$\ln \tilde{Y}_{x,t}$		Spread of $\ln \tilde{Y}_{x,t}$		Upper limit of $\ln \tilde{Y}_{x,t}$		Lower limit of $\ln \tilde{Y}_{x,t}$	
	Male	Female	Male	Female	Male	Female	Male	Female
2009	−7.97	−8.70	1.37	1.37	−6.61	−7.34	−9.34	−10.07
2010	−8.05	−8.79	1.38	1.38	−6.68	−7.41	−9.43	−10.17
2011	−8.13	−8.88	1.39	1.39	−6.74	−7.49	−9.53	−10.27
2012	−8.21	−8.96	1.41	1.41	−6.80	−7.56	−9.61	−10.37
2013	−8.28	−9.04	1.42	1.42	−6.86	−7.62	−9.70	−10.46
2014	−8.35	−9.12	1.43	1.43	−6.92	−7.68	−9.79	−10.55
2015	−8.42	−9.19	1.45	1.45	−6.98	−7.74	−9.87	−10.63
2016	−8.49	−9.26	1.46	1.46	−7.03	−7.80	−9.95	−10.72
2017	−8.55	−9.32	1.47	1.47	−7.08	−7.85	−10.03	−10.80
2018	−8.62	−9.38	1.49	1.49	−7.13	−7.90	−10.10	−10.87
2019	−8.68	−9.44	1.50	1.50	−7.18	−7.95	−10.18	−10.94
2020	−8.74	−9.50	1.51	1.51	−7.22	−7.99	−10.25	−11.01
2021	−8.79	−9.56	1.53	1.53	−7.27	−8.03	−10.32	−11.08
2022	−8.85	−9.61	1.54	1.54	−7.31	−8.07	−10.39	−11.15
2023	−8.90	−9.66	1.55	1.55	−7.35	−8.11	−10.45	−11.21
2024	−8.95	−9.71	1.57	1.57	−7.39	−8.14	−10.52	−11.27
2025	−9.00	−9.75	1.58	1.58	−7.42	−8.17	−10.58	−11.33
2026	−9.05	−9.80	1.59	1.59	−7.46	−8.20	−10.64	−11.39
2027	−9.10	−9.84	1.61	1.61	−7.49	−8.23	−10.71	−11.44
2028	−9.15	−9.88	1.62	1.62	−7.53	−8.26	−10.76	−11.50
2029	−9.19	−9.92	1.63	1.63	−7.56	−8.28	−10.82	−11.55
2030	−9.23	−9.95	1.65	1.65	−7.59	−8.31	−10.88	−11.60
2031	−9.28	−9.99	1.66	1.66	−7.62	−8.33	−10.93	−11.65
2032	−9.32	−10.02	1.67	1.67	−7.64	−8.35	−10.99	−11.69
2033	−9.36	−10.05	1.69	1.69	−7.67	−8.37	−11.04	−11.74
2034	−9.39	−10.08	1.70	1.70	−7.69	−8.38	−11.09	−11.78
2035	−9.43	−10.11	1.71	1.71	−7.72	−8.40	−11.14	−11.82
2036	−9.47	−10.14	1.73	1.73	−7.74	−8.41	−11.19	−11.86
2037	−9.50	−10.16	1.74	1.74	−7.76	−8.43	−11.24	−11.90
2038	−9.53	−10.19	1.75	1.75	−7.78	−8.44	−11.29	−11.94
2039	−9.57	−10.21	1.77	1.77	−7.80	−8.45	−11.33	−11.98
2040	−9.60	−10.24	1.78	1.78	−7.82	−8.46	−11.38	−12.02
2041	−9.63	−10.26	1.79	1.79	−7.84	−8.47	−11.42	−12.05
2042	−9.66	−10.28	1.81	1.81	−7.85	−8.47	−11.46	−12.08
2043	−9.69	−10.30	1.82	1.82	−7.87	−8.48	−11.51	−12.12
2044	−9.72	−10.32	1.83	1.83	−7.88	−8.49	−11.55	−12.15
2045	−9.74	−10.34	1.85	1.85	−7.90	−8.49	−11.59	−12.18
2046	−9.77	−10.35	1.86	1.86	−7.91	−8.50	−11.63	−12.21

(continued)

Table 3 (continued)

Year	$\ln \tilde{Y}_{x,t}$							
	$\ln \tilde{Y}_{x,t}$		Spread of $\ln \tilde{Y}_{x,t}$		Upper limit of $\ln \tilde{Y}_{x,t}$		Lower limit of $\ln \tilde{Y}_{x,t}$	
	Male	Female	Male	Female	Male	Female	Male	Female
2047	−9.79	−10.37	1.87	1.87	−7.92	−8.50	−11.67	−12.24
2048	−9.82	−10.39	1.89	1.89	−7.93	−8.50	−11.70	−12.27
2049	−9.84	−10.40	1.90	1.90	−7.94	−8.50	−11.74	−12.30
2050	−9.87	−10.42	1.91	1.91	−7.95	−8.51	−11.78	−12.33

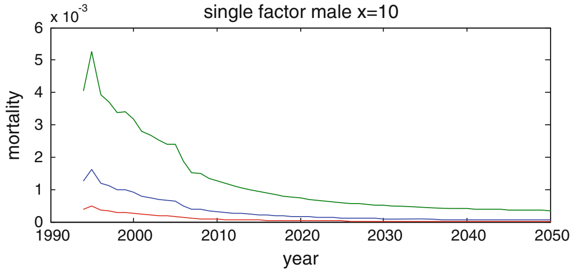


Fig. 6 Single factor male (x = 10, age group 27–29) mortality

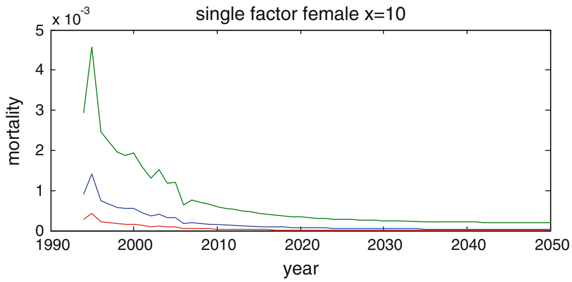


Fig. 7 Single factor female (x = 10, age group 27–29) mortality

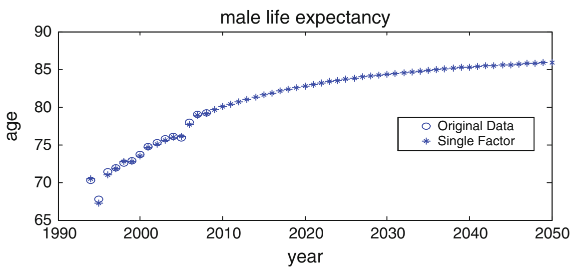


Fig. 8 Single factor male life expectancy

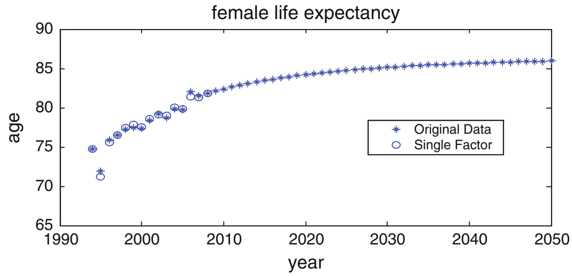


Fig. 9 Single factor female life expectancy

Table 4 $b_x^{(1)}, b_x^{(2)}$ with age-specific enhancement

x	Male		Female		x	Male		Female	
	$b_x^{(1)}$	$b_x^{(2)}$	$b_x^{(1)}$	$b_x^{(2)}$		$b_x^{(1)}$	$b_x^{(2)}$	$b_x^{(1)}$	$b_x^{(2)}$
1	0.0309	0.0710	0.0161	0.0359	16	0.0102	0.0077	0.0099	0.0154
2	0.0231	0.0466	0.0123	0.0163	17	0.0143	0.0159	0.0158	0.0280
3	0.0158	0.0209	0.0133	0.0139	18	0.0120	0.0070	0.0099	0.0116
4	0.0176	0.0257	0.0172	0.0272	19	0.0128	0.0127	0.0090	0.0134
5	0.0074	0.0037	0.0104	0.0133	20	0.0145	0.0105	0.0171	0.0132
6	0.0124	0.0213	0.0153	0.0300	21	0.0188	0.0160	0.0130	0.0138
7	0.0263	0.0617	0.0228	0.0536	22	0.0154	0.0091	0.0159	0.0113
8	0.0162	0.0323	0.0205	0.0475	23	0.0154	0.0085	0.0147	0.0143
9	0.0172	0.0214	0.0222	0.0556	24	0.0157	0.0091	0.0133	0.0134
10	0.0152	0.0277	0.0190	0.0394	25	0.0142	0.0138	0.0173	0.0109
11	0.0157	0.0218	0.0124	0.0248	26	0.0132	0.0131	0.0116	0.0124
12	0.0122	0.0136	0.0118	0.0123	27	0.0110	0.0112	0.0129	0.0111
13	0.0139	0.0256	0.0082	0.0104	28	0.0184	0.0166	0.0168	0.0084
14	0.0088	0.0034	0.0142	0.0240	29	0.0145	0.0150	0.0115	0.0076
15	0.0081	0.0070	0.0134	0.0176	30	0.0115	0.0062	0.0139	0.0087

(2) $k_t^{(2)'}$: We select ARIMA (0,2,1) for male and ARIMA (1,2,0) for female using ADF test and R test.

The results are as following (Figs. 10, 11, 12 and 13):

Table 5 $k_t^{(1)}, k_t^{(2)}$ and justified $k_t^{(2)'}$ with age-specific enhancement

T	Male			Female		
	$k_t^{(1)}$	$k_t^{(2)}$	$k_t^{(2)'}$	$k_t^{(1)}$	$k_t^{(2)}$	$k_t^{(2)'}$
1	10.4040	10.2330	11.8243	12.9460	12.1762	13.5574
2	17.5866	16.8844	16.9228	20.2478	20.1532	22.0503
3	9.1799	8.9434	9.4664	9.8251	8.5309	10.0529
4	8.9691	8.7685	8.4060	6.3601	6.7724	7.7928
5	6.1871	6.0817	5.4046	3.8698	4.2526	5.4819
6	5.5871	5.5057	6.0159	3.6555	3.3305	4.5185
7	2.5932	3.0934	3.2279	4.7060	4.0327	4.6582
8	−0.7181	−0.4087	0.6446	1.0441	1.3139	1.0722
9	−1.0705	−1.7893	−1.0068	−4.2277	−3.2679	−1.3828
10	−3.3719	−3.3982	−1.6451	−2.0950	−0.4707	−0.7302
11	−4.3293	−5.2782	−4.4506	−5.4636	−6.2049	−5.1482
12	−4.7509	−5.1788	−4.0797	−4.9691	−4.5718	−3.6425
13	−11.7264	−11.2937	−10.7726	−17.2602	−16.8087	−17.2865
14	−17.0810	−15.6848	−15.8321	−13.2665	−14.1426	−12.3293
15	−17.4589	−16.4782	−17.1257	−15.3723	−15.0958	−16.1282

Fig. 10 Male $k_t^{(1)}$ with age-specific enhancement

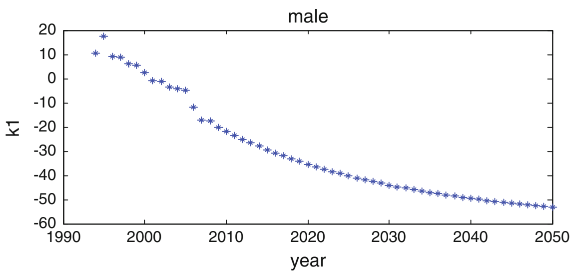
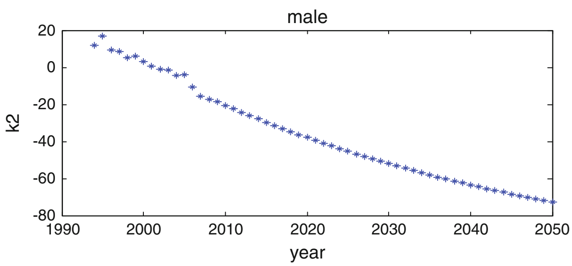


Fig. 11 Male $k_t^{(2)'}$ with age-specific enhancement



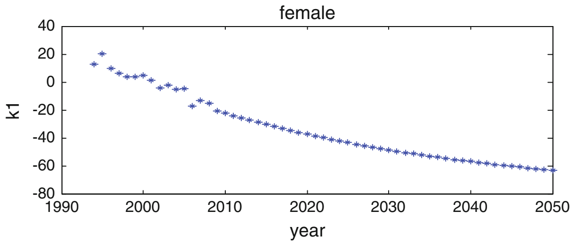


Fig. 12 Female $k_t^{(1)}$ with age-specific enhancement

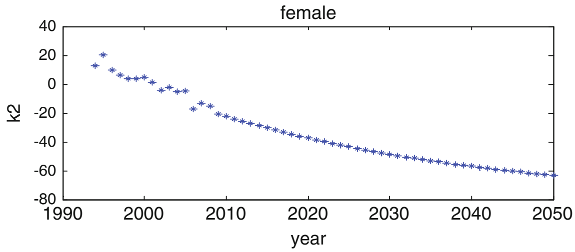


Fig. 13 Female $k_t^{(2)'}$ with age-specific enhancement

5 Conclusions and Remarks: China Data

Lastly, forecast the mortality from 2009 to 2050 with:

$$\hat{y}_{xt} = a_x + b_x^{(1)} k_t^{(1)} + b_x^{(2)} k_t^{(2)}, \quad t = t_0 + T, t_0 + T + 1, \ldots, t_0 + T + s$$

Then we can calculate the life expectancy and compare them with the original data (Figs. 14 and 15).

We summarize the results of life expectancy:

(1) In 2008, the expectancy of male is 79.36, the expectancy of female is 82.10.
(2) In 2050, the expectancy of male is 85.92, the expectancy of female is 86.86.
(3) The expectancy of male increased 6.56, the expectancy of female increased 4.76, and the difference of male and female decreased to 0.94 from 2.74.

We compared the fuzzy-LC with age-specific enhancement with single factor fuzzy-LC and found:

(1) The fuzzy-LC's sum error square with age-specific enhancement is smaller than the single factor fuzzy-LC's (Table 6).
(2) In both the fuzzy-LC's and fuzzy-LC's with age-specific enhancement, the expectancy of male and female increased from 2008 to 2050, and the difference of expectancy of male and female decreased from 2008 to 2050. The spread of mortality in Fuzzy-LC with age-specific enhancement.

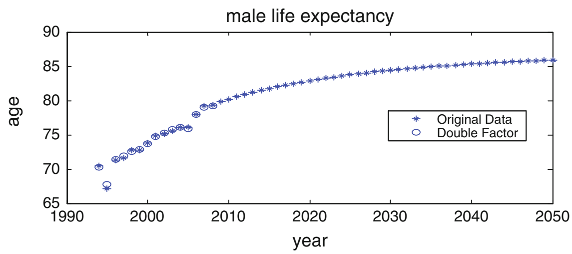


Fig. 14 Male life expectancy with age-specific enhancement

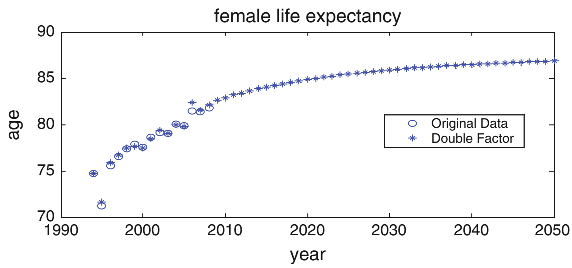


Fig. 15 Female life expectancy with age-specific enhancement

Table 6 The compare of the mortality sum error square between the fuzzy-LC's with age-specific enhancement and the single factor fuzzy-LC's

	Fuzzy LC's	Fuzzy LC's with age-specific enhancement	Proportion (%)
Male	3.07	3.02	98
Female	3.49	3.45	99

Based on the formula of spread:$e_{x,t} = s_{0x} + s_{1x} \times t$, t = 1, 2, 3, … 57.

We can calculate the spread of mortality (Table 7):

Then we can draw the mortality forecasting figure (Figs. 16 and 17):

In Figs. 16 and 17, the middle curve is the forecasting mortality, the upper and lower curves are the upper limit and lower limit of forecasting mortality.

Just as argued in Koissi and Shapiro [11] and other papers, the fuzzy approach is not limited by the homoscedasticity issue encountered in the SVD errors and it is an adequate alternative to the SVD method. Considering the scarce data of China, I think the fuzzy approach is more convenient and appropriate.

The first step is to solve the fuzzification of the log-central death rates based on minimum fuzziness criterion; the second step is to estimate the parameters of fuzzy-LC with age-specific enhancement. We achieve this task based on Diamond distance measure so that the actual total deaths and the total expected deaths for each t matches, and it is also convenient to compute the spreads of the parameters of the fuzzy-LC with age-specific enhancement.

As 89 is the highest life expectance quoted from the China Mortality Table, the fig of Chinese expected life up to 80 is relatively flat compared other countries.

Table 7 Male and female (x = 10) mortality forecasting with age-specific enhancement

	$\ln \tilde{Y}_{x,t}$							
Year	$\ln \tilde{Y}_{x,t}$		Spread of $\ln \tilde{Y}_{x,t}$		Upper limit of $\ln \tilde{Y}_{x,t}$		Lower limit of $\ln \tilde{Y}_{x,t}$	
	Male	Female	Male	Female	Male	Female	Male	Female
2009	−7.98	−8.84	1.37	1.37	−6.61	−7.47	−9.35	−10.20
2010	−8.06	−8.94	1.38	1.38	−6.68	−7.56	−9.44	−10.32
2011	−8.14	−9.03	1.39	1.39	−6.74	−7.64	−9.53	−10.42
2012	−8.21	−9.12	1.41	1.41	−6.81	−7.71	−9.62	−10.53
2013	−8.29	−9.21	1.42	1.42	−6.87	−7.79	−9.71	−10.63
2014	−8.36	−9.29	1.43	1.43	−6.93	−7.86	−9.79	−10.72
2015	−8.43	−9.37	1.45	1.45	−6.98	−7.92	−9.87	−10.82
2016	−8.50	−9.45	1.46	1.46	−7.04	−7.99	−9.96	−10.90
2017	−8.56	−9.52	1.47	1.47	−7.09	−8.05	−10.04	−10.99
2018	−8.63	−9.59	1.49	1.49	−7.14	−8.10	−10.11	−11.08
2019	−8.69	−9.66	1.50	1.50	−7.19	−8.16	−10.19	−11.16
2020	−8.75	−9.72	1.51	1.51	−7.24	−8.21	−10.26	−11.23
2021	−8.81	−9.78	1.53	1.53	−7.28	−8.26	−10.33	−11.31
2022	−8.87	−9.84	1.54	1.54	−7.33	−8.30	−10.40	−11.38
2023	−8.92	−9.90	1.55	1.55	−7.37	−8.35	−10.47	−11.45
2024	−8.98	−9.96	1.57	1.57	−7.41	−8.39	−10.54	−11.52
2025	−9.03	−10.01	1.58	1.58	−7.45	−8.43	−10.61	−11.59
2026	−9.08	−10.06	1.59	1.59	−7.49	−8.47	−10.67	−11.65
2027	−9.13	−10.11	1.61	1.61	−7.52	−8.50	−10.74	−11.72
2028	−9.18	−10.16	1.62	1.62	−7.56	−8.54	−10.80	−11.78
2029	−9.23	−10.20	1.63	1.63	−7.59	−8.57	−10.86	−11.84
2030	−9.27	−10.25	1.65	1.65	−7.63	−8.60	−10.92	−11.89
2031	−9.32	−10.29	1.66	1.66	−7.66	−8.63	−10.98	−11.95
2032	−9.36	−10.33	1.67	1.67	−7.69	−8.66	−11.03	−12.00
2033	−9.40	−10.37	1.69	1.69	−7.72	−8.68	−11.09	−12.05
2034	−9.44	−10.40	1.70	1.70	−7.75	−8.71	−11.14	−12.10
2035	−9.49	−10.44	1.71	1.71	−7.77	−8.73	−11.20	−12.15
2036	−9.52	−10.47	1.73	1.73	−7.80	−8.75	−11.25	−12.20
2037	−9.56	−10.51	1.74	1.74	−7.82	−8.77	−11.30	−12.25
2038	−9.60	−10.54	1.75	1.75	−7.85	−8.79	−11.35	−12.29
2039	−9.64	−10.57	1.77	1.77	−7.87	−8.80	−11.40	−12.34
2040	−9.67	−10.60	1.78	1.78	−7.89	−8.82	−11.45	−12.38
2041	−9.71	−10.63	1.79	1.79	−7.92	−8.84	−11.50	−12.42
2042	−9.74	−10.66	1.81	1.81	−7.94	−8.85	−11.55	−12.46
2043	−9.77	−10.68	1.82	1.82	−7.96	−8.86	−11.59	−12.50
2044	−9.81	−10.71	1.83	1.83	−7.97	−8.87	−11.64	−12.54
2045	−9.84	−10.73	1.85	1.85	−7.99	−8.89	−11.68	−12.58

(continued)

Table 7 (continued)

Year	$\ln \tilde{Y}_{x,t}$							
	$\ln \tilde{Y}_{x,t}$		Spread of $\ln \tilde{Y}_{x,t}$		Upper limit of $\ln \tilde{Y}_{x,t}$		Lower limit of $\ln \tilde{Y}_{x,t}$	
	Male	Female	Male	Female	Male	Female	Male	Female
2046	−9.87	−10.75	1.86	1.86	−8.01	−8.90	−11.73	−12.61
2047	−9.90	−10.78	1.87	1.87	−8.03	−8.91	−11.77	−12.65
2048	−9.93	−10.80	1.89	1.89	−8.04	−8.91	−11.81	−12.68
2049	−9.96	−10.82	1.90	1.90	−8.06	−8.92	−11.85	−12.72
2050	−9.98	−10.84	1.91	1.91	−8.07	−8.93	−11.89	−12.75

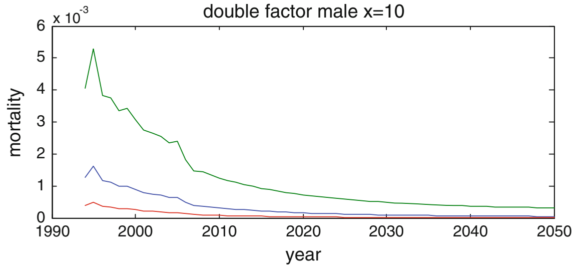


Fig. 16 Male (x = 10, age group 27–29) mortality curve with age-specific enhancement

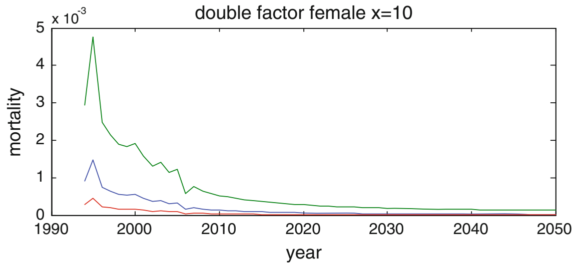


Fig. 17 Female (x = 10, age group 27–29) mortality curve with age-specific enhancement

References

1. Baser, F., Apaydin, A.: Calculating insurance claim reserves with hybrid fuzzy least squares regression analysis. Gazi Univ. J. Sci. **23**(10), 163–170 (2010)
2. Booth, H., Chauhan, R.K., Maindonald, J., Smith, L.: The future of the Australian mortality transition. In: The 2000 Biennial Conference of the Australian Population Association, Melbourne (2000)

3. Booth, H., Maindonald, J., Smith, L.: Age—time Interactions in Mortality Projection: Applying Lee. Carter to Australia. Working Papers in Demography, Australian National University, Canberra (2002)
4. Cummins, J.D., Derrig, R.A.: Fuzzy trends in property-liability insurance claim costs. J. Risk Insur. **60**(3), 429–465 (1993)
5. de Andres Sánchez, J.: Calculating insurance claim reserves with fuzzy regression. Fuzzy Sets Syst. **157**(23), 2091–3108 (2006)
6. Derrig, R.A., Ostaszewski, K.M.: Fuzzy techniques of pattern recognition in risk and claim classification. J. Risk Insur. **62**(3), 447–482 (1995)
7. Dewit, G.W.: Underwriting and uncertainty. Insur.: Math. Econ. **1**, 277–285 (1982)
8. Diamond, P.: Fuzzy least-squares. Inf. Sci. **46**(3), 141–157 (1988)
9. Ebanks, B., Kanvowski, W., Ostaszewski, K.M.: Application of measures of fuzziness to risk classification in insurance. In: Computing and Information. IEEE Computer Society Press, Los Alamitos (1992), pp. 290–291
10. Koissi, M.-C., Shapiro, A.F.: Fuzzy formulation of the Lee-Carter model for mortality forecasting. Insur.: Math. Econ. **39**, 287–309 (2006)
11. Lee, R.D., Carter, L.R.: Modeling and forecasting US mortality. J. Am. Stat. Assoc. **87**(419), 659–671 (1992)
12. Lemaire, J.: Fuzzy insurance. ASTIN Bulletin **20**(1), 33–55 (1990)
13. Ostaszewski, K.M.: Fuzzy Set Methods in Actuarial Science. Society of Actuarias, Schaumburg (1993)
14. Peizhuang, W.: Fuzzy Sets and Its Application. Shanghai Scientist Press, Beijing (1983)
15. Renshaw, A.E., Haberman, S.: A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insur.: Math. Econ., 556–570 (2006)
16. Renshaw, A.E., Haberman, S.: Lee-Carter mortality forecasting: a parallel generalized linear modeling approach for England and Wales mortality projections. Appl. Stat. **1**, **52**, 119–137 (2003)
17. Renshaw, A.E., Haberman, S.: On the forecasting of mortality reduction factors. Insur.: Math. Econ. **1**, **32**, 379–401 (2003)
18. Renshaw, A.E., Haberman, S.: Lee-Carter mortality forecasting with age—specific enhancement. Insur.: Math. Econ. **1**, **33**, 225–272 (2003)
19. Sánchez, J., Gómez, A.: Applications of regression in actuarial analysis. J. Risk Insur. **70**(4), 665–699 (2003)
20. Shapiro, A.F.: Fuzzy random variables. Insur.: Math. Econ. **44**, 307–314 (2009)
21. Shapiro, A.F.: Fuzzy logic in insurance. Insur.: Math. Econ. **35**, 399–424 (2004)
22. Young, V.R.: Insurance rate changing: a fuzzy logic approach. J. Risk Insur. **63**(3), 461–484 (1996)
23. Zimmermann, H.: Fuzzy Set Theory and its Application, 3rd edn. Kluwer Academic Publisher, Boston (1996)

Interval Number Comparison and Decision Making Based on Priority Degree

Yun-dong Gu, Su-jie Zhang and Meng-meng Zhang

Abstract Interval number ranking and the operation reduction are studied in this paper. Firstly, the priority relation of interval numbers is redefined by introducing a new priority-degree. We proved that the new priority-degree can avoid the unreasonable phenomenon that the denominator is zero when interval numbers degrade into real numbers. Then, some important comparison related concepts of the interval numbers, such as totally greater than, greater than, exactly equal and equal to etc., are redefined based on the new priority-degree. We also prove that interval numbers are reducible for addition and subtraction according to the redefined equal to relation. Finally, priority-degree based multi-criteria decision making method for uncertain problems with interval data is given and its validation is shown by an example.

Keywords Ranking interval number · Priority-degree · Equal to · Operation reduction

1 Introduction

Due to the incomplete information resulting from complexity and uncertainty in the decision making environments or lack of appropriate measurement units and scales, in some cases, it is difficult for decision makers to clearly translate their evaluation of alternatives in the form of exact numerical values. The representation of data by means of intervals of values is becoming more and more popular in different fields of application. Multiple criteria decision making (for short, MCDM) with interval numbers has been widely studied in many fields [1–19].

To deal with decision making problems in which alternative evaluation values or criterion weights are interval numbers, interval number ranking is one of the

Y. Gu · S. Zhang (✉) · M. Zhang
School of Mathematics and Physics, North China Electric Power University,
Beijing 102206, China
e-mail: zhangsujie12@126.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_19

key problems must be solved. For example, the concept of optimum or best choice to come true is completely based on ranking or comparison. For interval number ranking, many approaches have been proposed, such as the analytic hierarchy process, possibility degree, rough sets, probabilistic approach, etc. [20, 21]. In all of these approaches, one of the most commonly used methods is based on the concept of possibility degree, which is proposed by [2, 4, 5, 15, 22, 23], etc. However, when the interval numbers are degraded into real numbers, the existing formulas of possibility degree will appear the unreasonable phenomenon that the denominator will becoming zero when the interval numbers are degenerated into real numbers. To overcome this defection, a new priority-degree is introduced in this paper.

For interval number comparison, some relations such as greater than, no less than and equal to etc. are defined. Usually, interval numbers are not completely accurate under the uncertainty environment for lacking enough information, so we doubt that whether it is proper to set exact equal to relation for uncertainty interval numbers. On the contrary, we believe that a fuzzy equal to is more preferable due to the reason that it is much easier to operate and keeps more information for the following analysis. According to this idea, some interval number comparison related concepts, such as greater than and equal to etc., are redefined in this paper. It is worth observing that the operation reduction of interval numbers for addition and subtraction is proved for the redefined equal to relation.

The rest of this paper is organized as follows. A new priority-degree and its properties are given in Sect. 2. Interval number comparison related concepts are discussed in Sect. 3. Based on the redefined equal to relation, the reduction of interval numbers for addition and subtraction are proved in Sect. 4. In Sect. 5, a method for multi-criteria decision making based on priority degree is introduced and its validation is examined by an example. In Sect. 6, we conclude the paper.

2 Priority Degree and Interval Number Ranking

While ranking the alternative evaluations with uncertainty information in the form of interval numbers, the comparison of interval numbers is one of the most important basic problems. Quantitative indices are usually used to compare interval numbers (see reviews in [3, 7, 24]). One of the most commonly used methods is based on the concept of possibility degree.

2.1 Existing Possibility Degrees and Their Deficiencies

Interval numbers are a generalization of real numbers and are subsets of R, the set of real numbers. Normally, an interval number A is defined as

$$A = [a^L, a^R] = \{x | a^L < x < a^R, x \in R\} \tag{1}$$

where $a^L, a^R \in R$ are the lower and upper bounds of the interval number, respectively. Interval number A degenerates to a real number when $a^L = a^R$. Let I denote the set of interval numbers on R.

Definition 1 Let $A = [a^L, a^R] \in I$, then we say that A is a bounded closed interval number, if $-\infty < a^L \le a^R < +\infty$; we say that A is a positive bounded closed interval number, if $0 < a^L \le a^R < +\infty$.

Wang et al. [20] proposed a simple heuristic method which provides the degree of possibility that one interval is greater than another. Currently, this method seems to be most popular in the literature. For $A = [a^L, a^R], B = [b^L, b^R] \in I$, the possibility of $A \ge B$ given by [20] are defined as follows:

$$\mathrm{P}_1(A \ge B) = \frac{max(0, a^R - b^L) - max(a^L - b^R, 0)}{a^R + b^R - b^L - a^L} \tag{2}$$

The similar expressions were proposed by Facchinetti et al. [1] and by Xu and Da [5], Xu and Jain [25] showed that there are equivalent ones. [2, 4, 8, 14, 22, 23] etc., also introduced different possibilities, we list them as follows:

$$\mathrm{P}_2(A \ge B) = \frac{max\{0, a^R - a^L + b^R - b^L - max(b^R - a^L, 0)\}}{a^R + b^R - b^L - a^L} \tag{3}$$

$$\mathrm{P}_3(A \ge B) = \begin{cases} 1, b^R < a^L \\ \frac{a^R - b^L}{a^R - a^L + b^R - b^L}, b^L < a^R, a^L < b^R \\ 0, b^L \ll a^R \end{cases} \tag{4}$$

$$\mathrm{P}_4(A \ge B) = max\{1 - max(\frac{b^R - a^L}{a^R - a^L + b^R - b^L}, 0), 0\} \tag{5}$$

$$\mathrm{P}_5(A \ge B) = \begin{cases} 1, b^R < a^L \\ \frac{(a^R - b^L)^2}{(a^R - b^L)^2 + (b^R - a^L)^2}, b^L < a^R, a^L < b^R \\ 0, b^L \ge a^R \end{cases} \tag{6}$$

$$\mathrm{P}_6(A \ge B) = \frac{1}{2}(1 + \frac{(a^R - b^R) + (a^L - b^L)}{|a^R - b^R| + |a^L - b^L| + l_{AB}}) \tag{7}$$

where l_{AB} indicates the length of the overlap portion of two intervals.

$$\mathrm{P}_7(A \geq B) = \begin{cases} 1, b^R < a^L \\ 1 - \dfrac{(b^L - a^R)^2}{2l_A l_B}, b^L < a^L < b^R < a^R \\ \dfrac{a^L + a^R - 2b^L}{2l_B}, b^L < a^L < a^R < b^R \\ \dfrac{2a^R - b^L - b^R}{2l_B}, a^L < b^L < b^R < a^R \\ \dfrac{(a^R - b^L)^2}{2l_A l_B}, a^L < b^L < a^R < b^R \\ 0, a^R < b^L \end{cases} \tag{8}$$

Gao [15] had proved that the above possibility degrees $P_1 \sim P_4$ are mutual equal. It is easy to found that when the interval numbers are degenerated into real numbers, $P_1 \sim P_4$ and P_7 will appear the unreasonable phenomenon that the denominator is zero. And when the two comparing interval numbers are degraded into the same real numbers, all of the existing possibility degree formulas will appear the phenomenon of zero over zero. To overcome this imperfection, we introduce a new priority-degree.

2.2 Priority Degree and the Interval Number Ranking

First, we introduce the definition of priority degree.

Definition 2 Let $A = [a^L, a^R], B = [b^L, b^R] \in I$, then the priority degree between A and B can be given as

$$\mathrm{P}(A \geq B) = \begin{cases} 1, b^R < a^L \\ \dfrac{1}{2} + \dfrac{a^R - b^L + a^L - b^R}{2(a^R - a^L + b^R - b^L) + \varepsilon}, b^L < a^R, a^L < b^R \\ 0, b^L > a^R \end{cases} \tag{9}$$

where ε is a small positive parameter which is introduced to prevent the imperfection that denominator will become zero when interval numbers degenerates into real numbers. Usually we set $\varepsilon = 10^{-3}$ in calculation.

Ex. 1 Let $A = B = [a, a]$, by Definition 2 we get $P(A \geq B) = \frac{1}{2} + \frac{0}{2*0+\varepsilon} = \frac{1}{2}$.

Ex. 1 shows that the priority degree is free the imperfection for the known possibility degree that denominator will become zero when interval numbers degenerates into real numbers. In addition, it is easy to prove that it holds the following properties.

Proposition 1 *Let* $A = [a^L, a^R], B = [b^L, b^R], C = [c^L, c^R]$ *be three interval numbers, P is the priority-degree, then it holds that*

(a) $0 \leq P(A \geq B) \leq 1$.
(b) $P(A \geq A) = \frac{1}{2}$.
(c) $P(A \geq B) = 1 \Leftrightarrow a^L \geq b^R, P(A \geq B) = 0 \Leftrightarrow a^R \leq b^L$.
(d) $P(A \geq B) + P(B \geq A) = 1$.
(e) $P(A \geq B) = \frac{1}{2} \Leftrightarrow P(B \geq A) = \frac{1}{2}$.
(f) $P(A \geq B) \geq \frac{1}{2}, P(B \geq C) \geq \frac{1}{2} \Rightarrow P(A \geq C) \geq \frac{1}{2}$. $P(A \geq B) \leq \frac{1}{2}, P(B \geq C) \leq \frac{1}{2} \Rightarrow P(A \geq C) \leq \frac{1}{2}$. $P(A \geq B) = P(B \geq C) = \frac{1}{2} \Rightarrow P(A \geq C) = \frac{1}{2}$.

Proof On account of the properties are easily proved, so here we just prove one of them, the property (f), as an example. The other properties can be proved similarly.

$\mathrm{P}(A \geq B) \geq \frac{1}{2} \Leftrightarrow \frac{1}{2} + \frac{a^R - b^L + a^L - b^R}{2(a^R - a^L + b^R - b^L) + \varepsilon} \geq \frac{1}{2} \Leftrightarrow \frac{a^R - b^L + a^L - b^R}{2(a^R - a^L + b^R - b^L) + \varepsilon} \geq 0 \Leftrightarrow a^R - b^L + a^L - b^R \geq 0 \Leftrightarrow a^R + a^L \geq b^L + b^R$.

Similarly, we have $P(B \geq C) \geq \frac{1}{2} \Leftrightarrow b^R + b^L \geq c^L + c^R$. So it is obvious to prove that $a^R - c^L + a^L - c^R \geq 0$, and $P(A \geq C) = \frac{1}{2} + \frac{a^R - c^L + a^L - c^R}{2 * (a^R - a^L + cb^R - cb^L) + \varepsilon} \geq \frac{1}{2}$. So it holds that $P(A \geq B) \geq \frac{1}{2}, P(B \geq C) \geq \frac{1}{2} \Rightarrow P(A \geq C) \geq \frac{1}{2}$.

3 The Comparison of Interval Numbers

Based on the Definition 2 and the properties of priority degree, we can redefine some interval numbers comparing related concepts.

Definition 3 Let $A = [a^L, a^R], B = [b^L, b^R]$ are interval numbers, and P is the priority-degree. We say that A is totally greater than B or B is totally less than A, denotes as $A \succ B$ or $B \succ A$, if $P(A \geq B) = 1$. We say that A is greater than B or B is less than A, denotes as $A > B$ or $B > A$, if $P(A \geq B) > \frac{1}{2}$. We say that A is exactly equal to B, denotes as $A \equiv B$, if $a^L = b^L, a^R = b^R$. We say that A is equal to B, denotes as $A = B$, if $P(A \geq B) = P(B \geq A) = \frac{1}{2}$. We say that A is greater than B or equal to B or A is no less than B, denotes as $A \geq B$ or $B \leq A$, if $P(A \geq B) \geq \frac{1}{2}$. Especially, we say that A is non-negative if $A \geq 0$.

Based on Definition 3, we have the conclusion that

Lemma 1 *Let $A, B \in I$ and then it holds that $A \succ B \Rightarrow A > B, A \equiv B \Rightarrow A = B$ but not vice versa.*

Lemma 2 *Let $A, B \in I$ and then it holds that $A > B \Rightarrow A \succ B$, and $A = B \Rightarrow A \equiv$ Bholds if and only if A, B are degraded into real numbers.*

Lemma 3 *Suppose $A = [a^L, a^R]$ and $B = [b^L, b^R]$ are interval numbers, then it holds the following properties*

(a) $A \succ B \Leftrightarrow a^L \geq b^R$.
(b) $A > B \Leftrightarrow B < A \Leftrightarrow P(B \leq A) = 1$.
(c) $A \geq B \Leftrightarrow a^R + a^L \geq b^R + b^L$.
(d) $A = B \Leftrightarrow a^R + a^L = b^R + b^L, A = 0 = [0,0] \Leftrightarrow a^R + a^L = 0$.

Ex. 2 Let $A = [-1,2], B = [0,1], C = [1,3] \in I$, *then we have* $A = B, C > A$ *and* $C \succ B$.

4 The Reduction for Addition and Subtraction

The usually used arithmetic operations of interval numbers are defined as below:

(a) Addition: $[a,b] + [c,d] = [a+c, b+d]$.
(b) Subtraction: $[a,b] - [c,d] = [a-d, b-c]$.
(c) Multiplication: $[a,b] \times [c,d] = [min(ac, ad, bc, bd), max(ac, ad, bc, bd)]$.
(d) Division: $[a,b] \div [c,d] = [a,b] \times [\frac{1}{d}, \frac{1}{c}]$, where $[c,d]$ is positive bounded.

Lemma 4 *Suppose* $A = [a^L, a^R], B = [b^L, b^R] \in I$ *then it holds that* $A \geq B \Leftrightarrow A - B \geq 0 \Leftrightarrow B - A \leq 0$.

Proof By Definition 2, it holds that
$A \geq B \Leftrightarrow P(A \geq B) \geq \frac{1}{2} \Leftrightarrow \frac{a^R - b^L + a^L - b^R}{2(a^R - a^L + b^R - b^L) + \varepsilon} \geq 0 \Leftrightarrow \frac{a^R - b^L - 0^L + a^L - b^R - 0^R}{2((a^R - b^L) - (a^L - b^R) + 0^R - 0^L) + \varepsilon} \geq 0 \Leftrightarrow P(A - B \geq 0) \geq \frac{1}{2} \Leftrightarrow A - B \geq 0$.

So it holds that $A \geq B \Leftrightarrow A - B \geq 0$. The other part can be proved in the same way.

Lemma 5 *Suppose* $A = [a^L, a^R], B = [b^L, b^R] \in I$, *then it holds that* $A + B - B \equiv A \Leftrightarrow b^L = b^R$.

Proof Since $A + B - B = [a^L + b^L - b^R, a^R + b^R - b^L]$, by Definition 3, we have $A + B - B \equiv A \Leftrightarrow a^L + b^L - b^R = a^L, a^R + b^R - b^L = a^R \Leftrightarrow b^L = b^R$.

Lemma 5 means that according to the exactly equal relation, interval numbers are not reducible for addition and subtraction until the interval numbers are degrade to real numbers. This is the main reason that some references state that the interval numbers does not holds the reduction property for addition and subtraction. However, we argue that the reason of interval numbers is not reducible for addition and subtraction lies in the fact that the exactly equal to is too strict for the uncertainty numbers. Without enough information, soft concept may be better than precision one. In fact, according to the redefined equal to, the interval number are reducible for addition and subtraction operation.

Proposition 2 *Suppose that* $A = [a^L, a^R]$ *and* $BB = [b^L, b^R]$ *are two interval numbers, then it holds the following conclusion* $A + B - B = A = A - B + B$.

Proof Since $A + B - B = [a^L + b^L - b^R, a^R + b^R - b^L] = A - B + B$, it is easy to find that $a^L + b^L - b^R + a^R + b^R - b^L = a^L + a^R$. By Lemma 3, it holds that $A + B - B = A = A - B + B$.

Ex. 3 For $A = [-1, 2], B = [0, 1] \in I$, we have $A + B - B = A - B + B = [-2.3]$, $P(A + B - B \geq A) = \frac{1}{2}$, i.e., $A + B - B = A$.

Note: By redefining the equal relation of interval numbers, we discover the fact that although interval numbers do not reducible on addition and subtraction according exactly equal to or the same as relation, but it is reducible according to the equal relation defined in Definition 3. In other words, after the operation of adding an interval number and subtracting the same one, an interval number becomes another number in form, but it still is it in terms of value. The new one and the original are equal. The more operation, the more uncertainty the number is. But generally speaking, it is still unchanged in value. In fact, the exactly equal to is not suitable for the uncertain interval numbers.

5 Priority-Degree Based Multi-criteria Decision Making

By using the given priority degree measure and the concepts of interval number comparison, we can give a multi-criteria decision making methods, which can be described as following.

Priority-degree based multi-criteria decision making algorithm

Step 1: Considering factor set $F = f_1, f_2, \ldots, f_n$ to evaluate alternatives $A_1, A_2, \ldots, A_m$. Suppose the evaluation value of alternative A_i on factor f_j is described by interval numbers $A_{ij} = [a_{ij}^L, a_{ij}^R], i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$. The weight vector is $w = w_1, w_2, \ldots, w_n$.

Step 2: Compute the synthesized evaluation value of $F = f_1, f_2, \ldots, f_n$ by weighted average method, i.e. $A_i = \Sigma_{j=1}^n w_i A_{ij}, i = 1, 2, \ldots, m$.

Step 3: Let $\varepsilon = 10^{-4}$, compute the priority degree of A_i comparing with A_j, denote as $P_{ij}, i, j \in 1, 2, \ldots, n$, we get the priority degree matrix $P = (p_{ij})_{n \times n}$.

Step 4: Calculate the sorting vectors $v_i = \frac{1}{n(n-1)} \Sigma_{j=1}^n p_{ij} + \frac{n}{2} - 1, i = 1, 2, \ldots, n$.

Step 5: Rank the alternatives $A_1, A_2, \ldots, A_m$ by the value of $v_i, i = 1, 2, \ldots, n$.

In the following, we examine this method by an example. Example 5 Considering a college evaluation problems, adopt teaching, research and service as evaluation index, denoted by B_1, B_2, B_3. Table 1 lists the normalized evaluation data of five colleges $X_1, X_2, \ldots, X_5$. Suppose the weight vector is $w = (0.4, 0.4, 0.2)$.

Table 1: The evaluation results of five colleges

By the weighted average method, we get the synthesized evaluation values: $A_1 = [0.1888, 0.1972]$, $A_2 = [0.2068, 0.2198]$, $A_3 = [0.1988, 0.2070]$, $A_4 = [0.1874, 0.1970]$, $A_5 = [0.1874, 0.1962]$. Based on Definition 2, we compute the priority-degree matrix

Table 1 The evaluation results of five colleges

	X_1	X_2	X_3	X_4	X_5
B_1	[0.214, 0.220]	[0.206, 0.225]	[0.195, 0.204]	[0.181, 0.190]	[0.175, 0.184]
B_2	[0.166, 0.178]	[0.220, 0.229]	[0.192, 0.198]	[0.195, 0.205]	[0.193, 0.201]
B_3	[0.184, 0.190]	[0.182, 0.191]	[0.220, 0.231]	[0.185, 0.195]	[0.201, 0.211]

$$P = \begin{pmatrix} 0.5000 & 0.0000 & 0.0000 & 0.5451 & 0.5630 \\ 1.0000 & 0.5000 & 0.9906 & 1.0000 & 1.0000 \\ 1.0000 & 0.0094 & 0.5000 & 1.0000 & 1.0000 \\ 0.4549 & 0.0000 & 0.0000 & 0.5000 & 0.5224 \\ 0.4370 & 0.0000 & 0.0000 & 0.4776 & 0.5000 \end{pmatrix}. \tag{10}$$

Then, by formula (1), the sorting vector $V = (0.1554, 0.2995, 0.2505, 0.1489, 0.1457)$. It means that $X_2 \geq X_3 \geq X_1 \geq X_4 \geq X_5 3$.

6 Conclusion

In actual muti-criteria decision making process, sometimes it is difficult to determine precisely the real values of ratings of alternatives with respect to local criteria, and as a result, these ratings are presented by interval numbers. In this paper, a new priority-degree is introduced and the comparison related concepts are redefined based on it. It is worth to note that we prove the reduction of interval numbers for addition and subtraction. A multi-criteria decision making method based on given priority degree is given and its validation is shown by an example. In future, we will consider the applications of interval number comparison in the fields of portfolio selection, risk analysis, environment healthy or electric power quality evaluation etc.

Acknowledgments This paper is supported by the Fundamental Research Funds for the Central Universities (2015MS54), the Co-construction Project of Beijing Municipal commission of Education and the National Natural Science Foundation of China (Grant No. 71171080)

References

1. Facchinetti, G., Ricci, R.G., Muzzioli, S.: Note on ranking fuzzy triangular numbers. Int. J. Intell. Syst. **13**, 613–622 (1998)
2. Da, Q., Liu, X.: Interval number linear programming and its satisfactory solution. Syst. Eng. Theory Pract. **19**(4), 3–7 (1999)
3. Wang, X., Kerre, E.E.: Reasonable properties for the ordering of fuzzy quantities (I) (II). Fuzzy Sets Syst. **112**, 387–405 (2001)
4. Gao, F., Luo, Y.: Interval assignment problem research and application. Decision Science-Theory and Method, pp. 267–271. China Ocean Press, Beijing (2001)

5. Xu, Z., Da, Q.: The uncertain OWA operator. Int. J. Intell. Syst. **17**, 569C575 (2002)
6. Qi, Z., Zhang, W., Fan, Y.: A new multiple attribute decision making approach based on interval number. Oper. Res. Manage. Sci. **15**(4), 51–55 (2006)
7. Sevastianov, P.: Numerical methods for interval and fuzzy number comparison based on the probabilistic approach and dempster-shafer theory. Inf. Sci. **177**(21), 4645–4661 (2007)
8. Xu, Z., Da, Q.: Possibility degree method for ranking interval numbers and its application. J. Syst. Eng. **18**(1), 67–70 (2003)
9. Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Davoodi, A.R.: Extension of TOPSIS for decision-making problems with interval data: Interval efficiency. Math. Comput. Model. **49**(5–6), 1137–1142 (2009)
10. Sayadi, M.K., Heydari, M., Shahanaghi, K.: Extension of VIKOR method for decision making problem with interval numbers. Appl. Math. Model. **33**(5), 2257–2262 (2009)
11. Peng, A., Xiao, X.: Normalization methods for attribute values in fuzzy Multi-attribute decision making with interval numbers. Mach. Design Res. **27**(6), 5–8 (2011)
12. Huang, Z., Liu, J., Liu, S., et al.: Prospect theory model for multiple criteria decision making alternative with interval number. Syst. Eng. Electron. **34**(5), 977–981 (2012)
13. Zhang, Q., Fan, Z., Pan, D., et al.: A Ranking Approach for Interval Numbers in Uncertain Multiple Attribute Decision Making Problems. **19**(5), 129–133 (1999)
14. Xie, N., Liu, S.: On comparing grey numbers with their probability distribution. Syst. Eng.-Theory Practice **29**(4), 169–175 (2009)
15. Gao, F.: Possibility degree and comprehensive priority of interval numbers. Syst. Eng.-Theory Practice **33**(8), 2033–2040 (2013)
16. Wu, J., Jiasen, S., Malin, S., Liang, L.: A ranking method for DMUS with interval data based on DEA cross-efficiency evaluation and TOPSIS. J. Syst. Sci. Syst. Eng. **22**(2), 191–201 (2013)
17. Yue, Z.: Group decision making with multi-attribute interval data. Inform. Fusion **14**(4), 551–561 (2013)
18. Yang, Y., Li, J., Ding, W., et al.: Information flowing pattern construction for air offensive campaign based on interval number. Syst. Eng. Electron. **36**(9), 1761–1768 (2014)
19. Chang, Z., Cheng, L., Liu, J.: Multiple attribute decision making method with interval based on Mahalanobis-taguchi system and TOPSIS method. Syst. Eng.-Theory Practice. **34**(1), 168–175 (2014)
20. Wang, Y.-M., Yang, J.-B., Dong-Ling, X.: A two-stage logarithmic goal programming method for generating weights from interval comparison matrices. Fuzzy Sets Syst. **152**, 475–498 (2005)
21. Sevastjanov, P.: Numerical methods for interval and fuzzy number comparison based on the probabilistic approach and Dempster-Shafer theory. Inform. Sci. **177**, 4645C4661 (2007)
22. Lan, J., Cao, L., Lin, J.: Method for ranking interval numbers based on two-dimensional priority degree. J. Chongqing Inst. Technol. (Natural Science Edition) **21**(10), 63–67 (2007)
23. Li, D., Gu, Y.: Method for ranking interval numbers based on possibility degree. J. Syst. Eng. **23**(2), 223–226 (2008)
24. Xuesheng, L., Wei, W., Kaiqi, Z.: Rough sets ranking methodology for interval numbers. J. Dalian Univ. Technol. **48**(1), 143–146 (2008)
25. Xu, Z., Jian, C.: Some models for deriving the priority weights from interval fuzzy preference relations. Eur. J. Oper. Res. **184**, 266–280 (2008)

T-Absolute Truth Degree Theory of Formulas in Three-Valued Łukasiewicz Propositional Logic System

He Jin-Rui, Hui Xiao-Jing and Shuang Jing-Ning

Abstract In three-valued Łukasiewicz propositional logic system, T-absolute truth degree was proposed by combining absolute truth degree with T-truth degree. Also the T-absolute similarity degree and pseudo-distance between two formulas were defined by using the T-absolute truth degree, and some of properties were discussed. It was proved that operations such as ¬, →, ∨, ∧ were continuous in the T-absolute pseudo-distance space.

Keywords Three-valued Łukasiewicz logic system · T-absolute truth degree · T-absolute similarity degree · Pseudo-distance

1 Introduction

At the beginning of the 21st century, Professor Wang first introduced truth degree of formulas in the classical propositional logic, defined the similarity degree and pseudo-distance between two formulas by using the truth degree, established logic metric space and provided a possible framework for models of approximate reasoning. Therefore, a large number of scholars have launched research on truth degree. The literatures [1–11] introduced the numerical calculation into the mathematical logic, put forward the thought of degree reasoning and established the connection between the symbolic and numerical calculation. In the literature [12], a concept of absolute truth degree of the latter in n-valued Łukasiewicz propositional logic was introduced when the infinite product measure [13] in uniformly distributed probability spaces was evaded. Then an arbitrary formula of absolute truth degree could use computer calculating within finite steps. The literature [14] obtained the T-truth degree of formulas in two-valued propositional logic system

H. Jin-Rui (✉) · H. Xiao-Jing · S. Jing-Ning
College of Mathematics and Computer Science, Yan'an University,
Yan'an 716000, China
e-mail: 1223607029@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_20

and discussed its properties in logic operations. The T-similarity degree and T-pseudo metric of formulas were defined.

This paper obtained the T-absolute truth degree of formulas by combining absolute truth degree and T-truth degree, defined T-absolute similarity degree and pseudo-distance between formulas in three-valued Łukasiewicz propositional logic system. It was proved that operations such as ¬, →, ∨, ∧ were continuous in the T-absolute pseudo-distance space.

2 T-Absolute Truth Degree Theory of Formulas

Suppose S is the set of atomic formulas of $p_1, p_2, \ldots, p_n, F(S)$ is the free algebra of type $(\neg, \rightarrow)$ generated by S and the elements of $F(S)$ are formulas.

Definition 1 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n, A \in F(S)$, and $\bar{A}_T$ is the function induced by A which is limited on T. Then $\tau_T(A) = \frac{1}{3^n}\left|\overline{A_T}^{-1}(1)\right|$ is called the T-absolute truth degree of A. When $T = \phi, \tau_T(A) = 0$.

Because $\overline{A_T}^{-1}(1) = \overline{A}^{-1}(1) \cap T$, then $\tau_T(A) = \frac{1}{3^n}\left|\overline{A}^{-1}(1) \cap T\right|$. When $T = \{0, \frac{1}{2}, 1\}^n$ from the literature [12], $\tau_T(A) = \frac{1}{3^n}\left|\overline{A}^{-1}(1)\right|$, so see absolute truth degree as a special case of T-absolute truth degree. Clearly, $0 \leq \tau_T(A) \leq \frac{1}{3^n}|T|$.

Definition 2 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n$ and $A, B \in F(S)$. When $T = \Phi$ or when $T = \Phi$, if $\forall (x_1, x_2, \ldots x_n) \in T$, then $\bar{A}_T(x_1, x_2, \ldots, x_n) = \bar{B}_T(x_1, x_2, \ldots, x_n)$, A, B are T-logical equivalence, write $A \approx_T B$. A is T-absolute autology if $\tau_T(A) = \frac{1}{3^n}|T|$, and A is T-absolute contradiction if $\tau_T(A) = 0$.

Theorem 1 *Suppose* $T \subseteq \{0, \frac{1}{2}, 1\}^n$ *and* $A, B \in F(S)$, *then*:

$$\tau_T(A \vee B) = \tau_T(A) + \tau_T(B) - \tau_T(A \wedge B)$$

Proof Suppose A and B contain the same atomic formulas of $p_1, p_2, \ldots, p_n$,

From the literature [12], we can see $\overline{(A \vee B)}^{-1}(1) = \overline{A}^{-1}(1) \cup \overline{B}^{-1}(1)$,

So $\overline{(A \vee B)_T}^{-1}(1) = \overline{A_T}^{-1}(1) \cup \overline{B_T}^{-1}(1)$,

$$\begin{aligned}\left|\overline{A_T}^{-1}(1) \cup \overline{B_T}^{-1}(1)\right| &= \left|\overline{A_T}^{-1}(1)\right| + \left|\overline{B_T}^{-1}(1)\right| - \left|\overline{A_T}^{-1}(1) \cap \overline{B_T}^{-1}(1)\right| \\ &= \left|\overline{A_T}^{-1}(1)\right| + \left|\overline{B_T}^{-1}(1)\right| - \left|\overline{(A \wedge B)_T}^{-1}(1)\right|\end{aligned}$$

Therefore $\left|\overline{(A \vee B)_T}^{-1}(1)\right| = \left|\overline{A_T}^{-1}(1)\right| + \left|\overline{B_T}^{-1}(1)\right| - \left|\overline{(A \wedge B)_T}^{-1}(1)\right|$.

Divided by 3^n on both sides, then

$$\tau_T(A \vee B) = \tau_T(A) + \tau_T(B) - \tau_T(A \wedge B).$$

Theorem 2 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n, A, B, C \in F(S), 0 \le \alpha$ and $\beta \le 1$ then:

(1) (MP rule of T-absolute truth degree) If $\tau_T(A) \ge \alpha$ and $\tau_T(A \to B) \ge \beta$, then $\tau_T(B) \ge \alpha + \beta - \frac{1}{3^n}|T|$.
(2) (HS rule of T-absolute truth degree) If $\tau_T(A \to B) \ge \alpha$ and $\tau_T(B \to C) \ge \beta$, then $\tau_T(A \to C) \ge \alpha + \beta - \frac{1}{3^n}|T|$.
(3) If $| - (A \to B)_T$, then $\tau_T(A) \le \tau_T(B)$.

Proof suppose A and B contain the same atomic formulas of $p_1, p_2, \ldots, p_n$,

(1) Write

$$E_1 = \overline{A_T}^{-1}(1) = \{(x_1, x_2, \ldots, x_n) \in L_3^n | \overline{A_T}(x_1, x_2, \ldots, x_n) = 1\}$$
$$E_2 = \overline{(A \to B)_T}^{-1}(1) = \{(x_1, x_2, \ldots, x_n) \in L_3^n | \overline{A_T}(x_1, x_2, \ldots, x_n) \le \overline{B_T}(x_1, x_2, \ldots, x_n)\}$$
$$E_3 = \overline{B_T}^{-1}(1) = \{(x_1, x_2, \ldots, x_n) \in L_3^n | \overline{B_T}(x_1, x_2, \ldots, x_n) = 1\}$$

Clearly $E_1 \cap E_2 \subseteq E_3$ and $E_1 \cup E_2 \subseteq T$, then $|E_3| \ge |E_1 \cap E_2| = |E_1| + |E_2| - |E_1 \cup E_2| \ge |E_1| + |E_2| - |T|$,
Divided by 3^n on both sides, then

$$\tau_T(B) \ge \tau_T(A) + \tau_T(A \to B) - \frac{1}{3^n}|T| \ge \alpha + \beta - \frac{1}{3^n}|T|$$

(2) From the literature [12], we can see $| - (A \to B) \to ((B \to C) \to (A \to C))$
So $| - (A \to B)_T \to ((B \to C)_T \to (A \to C)_T)$ and $Q = (A \to B)_T \to ((B \to C)_T \to (A \to C)_T)$, then $\tau(Q) = \frac{1}{3^n}|T|$, using MP rule of T-truth degree twice can get the conclusion.
(3) Because $| - (A \to B)_T$, then $\tau_T(A \to B) = \frac{1}{3^n}|T|$, we can get the conclusion from (1) $\tau_T(B) \ge \tau_T(A) + \tau_T(A \to B) - \frac{1}{3^n}|T|$, hence $\tau_T(A) \le \tau_T(B)$.

Inference 1 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n$ and $A, B, C \in F(S)$, then

(1) if $\tau_T(A) = \tau_T(A \to B) = \frac{1}{3^n}|T|$, then $\tau_T(B) = \frac{1}{3^n}|T|$.
(2) if $\tau_T(A \to B) = \tau_T(B \to C) = \frac{1}{3^n}|T|$, then $\tau_T(A \to C) = \frac{1}{3^n}|T|$.

3 T-Absolute Similarity Degree and Pseudo-Distance Between Two Formulas

Based on the T-absolute truth degree of formula, give the T-absolute similarity degree between formulas in three-valued Łukasiewicz propositional logic system.

Definition 3 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n$ and $A, B \in F(S)$, define ξ_T by:

$$\xi_T(A,B) = \tau((A \to B)_T \wedge (B \to A)_T)$$

call $\xi_T(A,B)$ the T-absolute similarity degree between A and B.

Based on the T-absolute truth degree of formulas, the following theorem is established.

Theorem 3 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n$ and $A, B \in F(S)$, then:

(1) A and B are T-absolute truth degree, if only and if $A_T \approx B_T$.
(2) $\xi_T(A,A) = \frac{1}{3^n}|T|$.
(3) $\xi_T(A,B) = \xi_T(B,A)$.

Proof (1)–(3) are clearly established.

Definition 4 Suppose $T \subseteq \{0, \frac{1}{2}, 1\}^n$ and $A, B \in F(S)$. define ρ_T by:

$$\rho_T(A,B) = \frac{1}{3^n}|T| - \xi_T(A,B),$$

call $\rho_T(A,B)$ the T-absolute pseudo-distance between A and B, $(F(S), \rho_T)$ is called the T-absolute pseudo-distance space.

Proof suppose A, B and C contain the same atomic formulas of $p_1, p_2, \ldots, p_n$

$$\forall A, B, C \in F(S), \rho_T(A,A) = \frac{1}{3^n}|T| - \xi_T(A,A) = 0$$

$$\rho_T(A,B) = \frac{1}{3^n}|T| - \xi_T(B,A) = \rho_T(B,A)$$

$$\rho_T(A,B) + \rho_T(B,C) = \frac{1}{3^n}|T| - [\xi_T(A,B) + \xi_T(B,C) - \frac{1}{3^n}|T|] \geq \frac{1}{3^n}|T| - \xi_T(A,C) = \rho_T(A,C)$$

Hence ρ_T is the T-absolute pseudo-distance in F(S).

Theorem 4 In the T-absolute pseudo-distance space $(F(S), \rho_T)$, the unary logical connection $\neg$ and binary operation logical connections $\to$, $\vee$, $\wedge$ in pseudo-metric space are continuous.

Proof suppose A, B and C contain the same atomic formulas of $p_1, p_2, \ldots, p_n$

To prove operation logical connections $\neg$ is continuous, only to prove if $\lim_{n \to \infty} \rho_T(A_n, A) = 0$, then $\lim_{n \to \infty} \rho_T(\neg A_n, \neg A) = 0$.

Only to prove $\forall A, B \in F(S)$, there is $\rho_T(\neg A, \neg B) = \rho_T(A, B)$.

From the literature [12], we can see $(\neg A \to \neg B) \approx (B \to A)$ and $(\neg B \to \neg A) \approx (A \to B)$, so $(\neg A \to \neg B)_T \approx (B \to A)_T$ and $(\neg B \to \neg A)_T \approx (A \to B)_T$, then $(\neg A \to \neg B)_T \wedge (\neg B \to \neg A)_T \approx (B \to A)_T \wedge (A \to B)_T$,

Therefore $\tau_T((\neg A \to \neg B) \wedge (\neg B \to \neg A)) = \tau_T((B \to A) \wedge (A \to B))$

Though Definition 4, there is:

$$\begin{aligned}\rho_T(\neg A, \neg B) &= \frac{1}{3^n}|T| - \xi_T(\neg A, \neg B) = \frac{1}{3^n}|T| - \tau_T((\neg A \to \neg B) \wedge (\neg B \to \neg A)) \\ &= \frac{1}{3^n}|T| - \tau_T((B \to A) \wedge (A \to B)) = \frac{1}{3^n}|T| - \zeta_T(B, A) = \rho_T(B, A) = \rho_T(A, B)\end{aligned}$$

Prove the logical connections $\to$ is continuous next. That is to prove if $\lim_{n \to \infty} \rho_T(A_n, A) = 0$ and $\lim_{n \to \infty} \rho_T(B_n, B) = 0$, then $\lim_{n \to \infty} \rho_T(A_n \to B_n, A \to B) = 0$.

Only to prove the following established:

$$\rho_T(A_n \to B_n, A \to B) \le \rho_T(A_n, A) + \rho_T(B_n, B)$$

Though Definition 4, there is

$$\rho_T(A_n \to B_n, A \to B) \le \rho_T(A_n \to B_n, A_n \to B) + \rho_T(A_n \to B, A \to B)$$

Therefore only to prove

$$\rho_T(A_n \to B_n, A_n \to B) \le \rho_T(B_n, B), \rho_T(A_n \to B_n, A \to B) \le \rho_T(A_n, A)$$

The example $\rho_T(A_n \to B_n, A_n \to B) \le \rho_T(B_n, B)$ proves as follows:

From the literature [12], we can see $|-(B_n \to B) \to ((A_n \to B_n) \to (A_n \to B))$ and $|-(B \to B_n) \to ((A_n \to B) \to (A_n \to B_n))$ So $|-(B_n \to B)_T \to ((A_n \to B_n)_T \to (A_n \to B)_T)$ and $|-(B \to B_n)_T \to ((A_n \to B)_T \to (A_n \to B_n)_T)$ then $|-((B_n \to B)_T \wedge (B \to B_n)_T) \to ((A_n \to B_n)_T \to (A_n \to B)_T) \wedge ((A_n \to B)_T \to (A_n \to B_n)_T))$

Based on the (3) of Theorem 2:

$$\tau((B_n \to B)_T \wedge (B \to B_n)_T) \le \tau(((A_n \to B_n)_T \to (A_n \to B)_T) \wedge ((A_n \to B)_T \to (A_n \to B_n)_T))$$

Then $\xi_T(B_n, B) \le \xi_T(A_n \to B_n, A_n \to B)$

Therefore $\rho_T(B_n, B) \ge \rho_T(A_n \to B_n, A_n \to B)$.

Similarly $\rho_T(A_n, A) \ge \rho_T(A_n \to B, A \to B)$

Hence $\rho_T(A_n \to B_n, A \to B) \le \rho_T(A_n, A) + \rho_T(B_n, B)$. So binary operation logical connections $\to$ is continuous.

From the literature [1], we can see $A \vee B = (A \to B) \to B, A \wedge B = \neg(\neg A \vee \neg B)$ So $(A \vee B)_T = (A \to B)_T \to B_T, (A \wedge B)_T = \neg(\neg A \vee \neg B)_T$, hence $\wedge$ and $\vee$ are continuous.

4 Conclusion

In three-valued £ukasiewicz propositional logic system, it proposed T-absolute truth degree, T-absolute similarity degree, T-absolute pseudo-distance and discussed their properties.

T-absolute pseudo-distance space was built. T-absolute truth degree was a limited truth degree, so it still needed to further study based on approximate reasoning of T-absolute truth degree.

References

1. Wang, G.J.: Introduction to Mathematical Logic and Resolution Principle. Science in China Press, Beijing (2006)
2. Wang, G.J.: Nonclassical Mathematical Logic and Approximate. Science in China Press, Beijing (2008)
3. Adams, E.W.: A Primer of Probability Logi. CSLI Publications, Stanford (1998)
4. Baioletti, M., et al.: Simplification rules for the coherent probability assessment problem. Ann. Math. Artif. Intell. **35**(12), 11–28 (2002)
5. Wang, G.J., Fu, L., Song J.S.: Theory of Truth Degrees in classical logic system. Sci. China (Ser. A) **31**(11), 998–1008 (2001)
6. Wang, G.J., Leung, Y.: Integrated semantics and logic metric spaces. Fuzzy sets Syst. **136**(1), 71–91 (2003)
7. Wang, G.J., Zhang, W.X.: Consistency degrees of finite theories in Łukasiewicz propositional Fuzzy logic. Fuzzy Sets Syst. **149**(2), 275–284 (2005)
8. Zhou, X.N., Wang, G.J.: Consistency degrees of theories in some systems of propositional fuzzy logic. Fuzzy Sets Syst. **152**(2), 321–331 (2005)
9. Wang, G.J., Li, B.J.: Theory of truth degrees of formulas in Łukasiewicz n-valued propositional logic and a limit theorem. Sci. China (Ser. E), **35**(6), 561–569 (2005)
10. Wang, G.J., Hu, M.D.: Z(2)-normable linear structure on two-valued propositional logic meric space. Acta Electronica Sin. **39**(4), 899–905 (2011)
11. Li, J., Li, S.P., Xia, Y.F.: Theory of Truth Degrees in Łukasiewicz n-valued propositional logic. Acta Math. Sin. **47**(4), 769–780 (2004)
12. Li, J., Li, J.S., Zhou, Y.: Absolute truth degree theory formulas in n-valued Łukasiewicz propositional fuzzy logic. J. Lanzhou Univ. Technol. **34**(1), 134–138 (2008)
13. Halmos, P.R.: Measure Theory. Springer, New York (1980)
14. Wang, T.M.: T-truth degree theory in two-propositional logic system. Comput. Eng. Appl. **46** (10), 33–35 (2010)

Author Biographies

He Jin-Rui was born in 1989. She is a M.S. candidate at Yan'an university. Her research is non-classical logic. 1223607029@qq.com.

Hui Xiao-Jing was born in 1973. She received the Ph.D. degree from shaan xi normal university 2008. Now She is an associate professor and master supervisor at Yan'an University, and the member of CCF. Her research interest is non-classical logic. xhmxiaojing@163.com.

Shuang Jing-Ning was born in 1989. She is a M.S. candidate at Yan'an university. Her research is non-classical logic. 1264905987@qq.com.

Universal Function Projective Synchronization of Chaotic Systems with Uncertainty by Using Active Fuzzy Sliding Mode Control

Baojie Zhang

Abstract In the paper, the definition of universal function projective synchronization (UFPS) is restated which generalize the definition of modified function projective synchronization. An active fuzzy sliding mode controller is introduced for UFPS of chaotic systems. By using the fuzzy logical system, the chattering phenomenon is reduced. The systems with uncertainty are considered. Theoretical proofs are given based on the Lyapunov stability theory. Simulation results verify the effectiveness of the introduced approach.

Keywords Universal function projective synchronization (UFPS) · Active fuzzy sliding mode control · Uncertainty

1 Introduction

Since the pioneer work of Pecora and Carroll [1] was introduced in 1990, chaotic synchronization has been a hot research field and attracted extensive attention. Especially, it has potential applications in many aspects such as secure communication, digital communication, power electronic devices and power quality, biological systems, chemical analysis and design, and information processing, etc. Meanwhile many types of chaotic synchronization have been proposed, e.g. complete synchronization (CS) [1, 2], phase synchronization (PHS) [3], antisynchronization (AS) [4–6], projective synchronization (PS) [7, 8], modified projective synchronization

B. Zhang (✉)
School of Mathematics and Statistics, Qujing Normal University,
Qujing 655011, People's Republic of China
e-mail: qjupk123@mail.dlut.edu.cn

B. Zhang
Faculty of Electronic Information and Electrical Engineering,
Dalian University of Technology, Dalian 116024, People's Republic of China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_21

(MPS) [9, 10], function projective synchronization (FPS) [11–13], and modified function projective synchronization (MFPS) [14].

Recently, MFPS was extensively investigated [15, 16], in which a scaling function matrix is involved into the synchronization scheme between the drive and response systems. We believe sometimes there exists a "displacement" between the synchronized systems. In [17], we introduced a new type of chaotic synchronization—universal projective synchronization. It is a generalization of MFPS of chaotic systems. For clarity, we'd better call it universal function projective synchronization (UFPS). In the definition of UFPS, a reference system was introduced to denote the "displacement". UFPS is a generalization of CS, AS, PS, MPS, FPS, MFPS depending on the forms of reference system and scaling function matrix. Therefore, it is interesting to investigate UFPS of chaotic systems.

In reality, linear and nonlinear uncertainty even noise frequently appears in control systems. Sliding mode controller has been proven to be an efficient tool to control complex nonlinear dynamic plants operating under uncertainty conditions [18]. In the field of chaotic synchronization, active sliding mode control is a typical method to synchronize two identical or different chaotic systems [10, 19–25].

How to eliminate the chattering in sliding mode control is very important. Using fuzzy system is an efficient method to weaken chattering [26]. An active fuzzy sliding mode control scheme is given which can be used to accomplish UFPS of chaotic systems, in which the systems with uncertainty are considered.

The organization of this paper is as follows. In Sect. 2, the definition of two forms of UFPS is recalled. In Sect. 3, System description and problem formulation are given. In Sect. 4, an active fuzzy sliding mode controller is introduced. Simulation results are provided in Sect. 5. The conclusion is finally drawn in Sect. 6.

2 Definition of UFPS

The drive system and the response system are defined as follows

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t)), \tag{1}$$

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{g}(\boldsymbol{y}(t)) + \boldsymbol{u}(t). \tag{2}$$

Here, $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T, \boldsymbol{y} = (y_1, y_2, \dots, y_n)^T \in \boldsymbol{R}^n$ are the state vectors. $f, g : \boldsymbol{R}^n \to \boldsymbol{R}^n$ are continuous nonlinear vector function. $\boldsymbol{u}(t) \in \boldsymbol{R}^n$ is the controller to be designed. If there is no controller $\boldsymbol{u}(t)$, Eqs. (1) and (2) are chaotic systems.

Suppose there exists a reference system

$$\dot{z}(t) = \boldsymbol{h}(z(t)) \tag{3}$$

where $\boldsymbol{z} = (z_1, z_2, \dots, z_n)^T \in \boldsymbol{R}^n$, $\boldsymbol{h} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is a continuous differentiable function. Here the reference system is an attractor.

Let us denote

$$\boldsymbol{e}(t) = (e_1(t), e_2(t), \ldots, e_n(t))^T \triangleq \boldsymbol{x}(t) - \boldsymbol{A}(t)\boldsymbol{y}(t) - \boldsymbol{z}(t) \tag{4}$$

where $\boldsymbol{A}(t) = diag(\alpha_1(t), \alpha_2(t), \ldots, \alpha_n(t))$, $\alpha_i(t)(i = 1, 2, 3, \ldots, n)$ is continuous differentiable function with bounded, and $\alpha_i(t) \neq 0$ for all t. We call $\boldsymbol{A}(t)$ the scaling function matrix. Obviously, we have

$$\dot{\boldsymbol{e}} = \boldsymbol{f}(\boldsymbol{x}(t)) - \dot{\boldsymbol{A}}(t)\boldsymbol{y}(t) - \boldsymbol{A}(t)(\boldsymbol{g}(\boldsymbol{y}(t)) + \boldsymbol{u}(t)) - \boldsymbol{h}(\boldsymbol{z}(t)) \tag{5}$$

Definition 1 (*UFPS*) For the scaling diagonal matrix function $\boldsymbol{A}(t)$, it is said that the drive system (1) and the response system (2) are universal function projective synchronization (UFPS) in the sense of the system (3), if the system (5) is asymptotically stable for any initial value $\boldsymbol{x}(0), \boldsymbol{y}(0), \boldsymbol{z}(0)$, i.e.,

$$\lim_{t\to\infty} \|\boldsymbol{e}\| = 0 \tag{6}$$

where the error $\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{A}(t)\boldsymbol{y} - \boldsymbol{z}$.

Remark 1 The reference system (3) expresses a constant vector $\boldsymbol{\xi}$, if it's initial value is $\boldsymbol{\xi}$ and the state change rate is $\boldsymbol{0}$. Now we call UFPS in the sense of constant vector $\boldsymbol{\xi}$, i.e. there exist a "displacement" $\boldsymbol{\xi}$ between the drive system and response system. System (5) can be other attractors, such as period function, quasi-period function, chaos and hyperchaos.

3 System Description and Problem Formulation

The drive system (7), the response system (8) and the reference system (9) are defined below

$$\dot{\boldsymbol{x}} = (\boldsymbol{B}_1 + \Delta\boldsymbol{B}_1)\boldsymbol{x} + \boldsymbol{g}_1(\boldsymbol{x}) + \Delta\boldsymbol{g}_1(\boldsymbol{x}) \tag{7}$$

$$\dot{\boldsymbol{y}} = (\boldsymbol{B}_2 + \Delta\boldsymbol{B}_2)\boldsymbol{y} + \boldsymbol{g}_2(\boldsymbol{y}) + \Delta\boldsymbol{g}_2(\boldsymbol{y}) + \boldsymbol{u}(t) \tag{8}$$

$$\dot{\boldsymbol{z}} = (\boldsymbol{B}_3 + \Delta\boldsymbol{B}_3)\boldsymbol{z} + \boldsymbol{g}_3(\boldsymbol{z}) + \Delta\boldsymbol{g}_3(\boldsymbol{z}) \tag{9}$$

where $\boldsymbol{x}(t) \in \boldsymbol{R}^n$, $\boldsymbol{y}(t) \in \boldsymbol{R}^n$ and $\boldsymbol{z}(t) \in \boldsymbol{R}^n$ denote the state vectors of the systems. $\boldsymbol{B}_1, \boldsymbol{B}_2$ and $\boldsymbol{B}_3 \in \boldsymbol{R}^{n\times n}$ represent coefficient matrices of the linear parts in the dynamic systems. $\boldsymbol{g}_1 : \boldsymbol{R}^n \to \boldsymbol{R}^n$, $\boldsymbol{g}_2 : \boldsymbol{R}^n \to \boldsymbol{R}^n$ and $\boldsymbol{g}_3 : \boldsymbol{R}^n \to \boldsymbol{R}^n$ are the nonlinear parts of the systems. $\Delta\boldsymbol{B}_1 \in \boldsymbol{R}^{n\times n}$, $\Delta\boldsymbol{B}_2 \in \boldsymbol{R}^{n\times n}$ and $\Delta\boldsymbol{B}_3 \in \boldsymbol{R}^{n\times n}$ are coefficient matrices of unknown linear parts and they are bounded. $\Delta\boldsymbol{g}_1 : \boldsymbol{R}^n \to \boldsymbol{R}^n$, $\Delta\boldsymbol{g}_2 : \boldsymbol{R}^n \to \boldsymbol{R}^n$ and $\Delta\boldsymbol{g}_3 : \boldsymbol{R}^n \to \boldsymbol{R}^n$ are unknown nonlinear parts of the drive system, response system and response system, respectively. The synchronization goal is as follows:

$$\lim_{t\to\infty} \|\boldsymbol{e}\| = \lim_{t\to\infty} \|\boldsymbol{x}(t) - \boldsymbol{A}(t)\boldsymbol{y}(t) - \boldsymbol{z}(t)\| = 0 \tag{10}$$

From Eqs. (7)–(9), we know the error dynamic system is

$$\begin{aligned}\dot{\boldsymbol{e}} &= \boldsymbol{B}_1 x + \boldsymbol{g}_1(x) - \dot{\boldsymbol{A}}(t)\boldsymbol{y} - \boldsymbol{A}(t)(\boldsymbol{B}_2\boldsymbol{y} + \boldsymbol{g}_2(\boldsymbol{y}) + \boldsymbol{u}(t)) \\ &\quad - \boldsymbol{B}_3 z - \boldsymbol{g}_3(z) + \boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, z).\end{aligned} \tag{11}$$

where $\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, z) = \Delta\boldsymbol{B}_1 x + \Delta\boldsymbol{g}_1(\boldsymbol{x}) - \boldsymbol{A}(t)\Delta\boldsymbol{B}_2\boldsymbol{y} - \boldsymbol{A}(t)\Delta\boldsymbol{g}_2(\boldsymbol{y}) - \Delta\boldsymbol{B}_3 z - \Delta\boldsymbol{g}_3(z)$.

4 Active Fuzzy Sliding Mode Control

Inspired by [24, 26] and considering the weakening chattering, we presented an active fuzzy sliding mode controller for UFPS of chaotic systems. Let the control input

$$\begin{aligned}\boldsymbol{u} &= \boldsymbol{A}^{-1}(t)(\boldsymbol{g}_1(\boldsymbol{x}) - \dot{\boldsymbol{A}}(t)\boldsymbol{y} - \boldsymbol{A}(t)\boldsymbol{B}_2\boldsymbol{y} - \boldsymbol{A}(t)\boldsymbol{g}_2(\boldsymbol{y}) \\ &\quad - \boldsymbol{B}_3 z - \boldsymbol{g}_3(z) + \boldsymbol{B}_1\boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{B}_1 z - \boldsymbol{H}(t))\end{aligned} \tag{12}$$

where $\boldsymbol{H}(t) = \boldsymbol{K}\boldsymbol{w}_1(t)$ is to be designed based on sliding mode control.

Here $\boldsymbol{w}_1(t) \in \boldsymbol{R}^n$ is a given by

$$\boldsymbol{w}_1(t) = \begin{cases} \boldsymbol{w}_1^+(t), \boldsymbol{s}(\boldsymbol{e}) \geq 0 \\ \boldsymbol{w}_1^-(t), \boldsymbol{s}(\boldsymbol{e}) < 0 \end{cases} \tag{13}$$

where $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{e}) = (s_1, s_2, \ldots, s_n)^T$ is the switching surface to be designed.

Substituting (12) into (11), we get

$$\dot{\boldsymbol{e}} = \boldsymbol{K}\boldsymbol{w}_1(t) + \boldsymbol{B}_1\boldsymbol{e} + \boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, z) \tag{14}$$

Now we take sliding mode surface

$$\boldsymbol{s}(\boldsymbol{e}) = \boldsymbol{C}\boldsymbol{e} \tag{15}$$

where C is a diagonal matrix, the elements on the diagonal are all positive constants.

To make the trajectory of the system lie in the sliding mode surface, two conditions

$$\boldsymbol{s} = \boldsymbol{0} \tag{16}$$

and

$$\dot{\boldsymbol{s}} = \boldsymbol{0} \tag{17}$$

are required.

To design the sliding mode controller, we choose the reaching law as:

$$\dot{\boldsymbol{s}} = -\boldsymbol{q}_1 sgn(\boldsymbol{s}) - r_1 \boldsymbol{s} \tag{18}$$

where $\boldsymbol{q}_1 = diag(q_{11}, q_{12}, \ldots, q_{1n})$ and r_1 is positive gain.

Considering Eqs. (15), (18), we get

$$\boldsymbol{w}_1(t) = -(\boldsymbol{CK})^{-1}(\boldsymbol{CB}_1\boldsymbol{e} + \boldsymbol{CM}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) + \boldsymbol{q}_1 sgn(\boldsymbol{s}) + r_1\boldsymbol{s}) \tag{19}$$

In practical engineering, the uncertainty $\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ are unknown. We expect to compensate uncertainty $\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ by increasing the value of $\boldsymbol{q}_1$. So the control input $\boldsymbol{w}_1(t)$ is rewritten as follows

$$\boldsymbol{w}_1(t) = -(\boldsymbol{CK})^{-1}(\boldsymbol{CB}_1\boldsymbol{e} + \boldsymbol{q}_1 sgn(\boldsymbol{s}) + r_1\boldsymbol{s}) \tag{20}$$

Suppose unknown nonlinear parts $\Delta\boldsymbol{g}_1(\boldsymbol{x})$, $\Delta\boldsymbol{g}_2(\boldsymbol{y})$ and $\Delta\boldsymbol{g}_3(z)$ are Lipschitz, i.e., $|\Delta\boldsymbol{g}_i(\boldsymbol{x}_1) - \Delta\boldsymbol{g}_i(\boldsymbol{x}_2)| \leq L_i \mid \boldsymbol{x}_1 - \boldsymbol{x}_2 \mid, \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{R}^n$ for some $L_i > 0$, where $|\cdot|$ denotes the norm of each component of a vector. And let $\Delta\boldsymbol{g}_1(\boldsymbol{0}) = \boldsymbol{0}$, $\Delta\boldsymbol{g}_2(\boldsymbol{0}) = \boldsymbol{0}$, $\Delta\boldsymbol{g}_3(\boldsymbol{0}) = \boldsymbol{0}$. We have:

$$|\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})| \leq N_1 \mid \boldsymbol{e}| + N_2 \mid \boldsymbol{y}| + N_3 \mid \boldsymbol{z} \mid \tag{21}$$

where $N_1 = \|\Delta\boldsymbol{B}_1\| + L_1$, $N_2 = \|\Delta\boldsymbol{B}_1\boldsymbol{A}(t) - \boldsymbol{A}(t)\Delta\boldsymbol{B}_2\| + (L_1 + L_2) \| \boldsymbol{A}(t) \|$ and $N_3 = \|\Delta\boldsymbol{B}_1 - \Delta\boldsymbol{B}_3\| + L_1 + L_3$. Let $N_1 \leq \alpha$, $N_2 \leq \beta$ for all $t \in [0, \infty)$ and $N_3 \leq \gamma$. Now we prove the stability of the combination of drive system, response system and the reference system under the designed controller (12).

Theorem 1 *The universal function projective synchronization of system (7) and system (8) in the sense of system (9) can occur by using the controller (12) with the switch part* $\boldsymbol{w}_1(t)$ *in Eq.* (20).

Proof Let us define the Lyapunov function as:

$$V = \frac{1}{2}\boldsymbol{s}^T\boldsymbol{s}. \tag{22}$$

The time derivative of Eq. (22) is:

$$\dot{V} = \boldsymbol{s}^T\dot{\boldsymbol{s}} \tag{23}$$

From Eqs. (14), (15) and (19), we have

$$\dot{\boldsymbol{s}} = -\boldsymbol{q}_1 sgn(\boldsymbol{s}) - r_1\boldsymbol{s} + \boldsymbol{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}). \tag{24}$$

Then

$$\begin{aligned}\dot{V} &= \boldsymbol{s}^T(-\boldsymbol{q}_1 sgn(\boldsymbol{s}) - r_1\boldsymbol{s} + \boldsymbol{M}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}))\\ &\le -\boldsymbol{q}_1\boldsymbol{s}^T sgn(\boldsymbol{s}) - r_1\boldsymbol{s}^T\boldsymbol{s} + |\boldsymbol{s}|^T(\alpha\mid \boldsymbol{e}\mid +\beta\mid \boldsymbol{y}\mid +\gamma\mid \boldsymbol{z}\mid)\\ &\le -\boldsymbol{q}_1\boldsymbol{s}^T sgn(\boldsymbol{s}) + |\boldsymbol{s}|^T(-r_1\mid \boldsymbol{s}\mid +\alpha\mid \boldsymbol{e}\mid +\beta\mid \boldsymbol{y}\mid +\gamma\mid \boldsymbol{z}\mid)\\ &< 0.\end{aligned} \tag{25}$$

$\dot{V}$ will be negative definite, if $\boldsymbol{s}$ meets the following condition:

$$\mid \boldsymbol{s}\mid > \frac{\alpha\mid \boldsymbol{e}\mid +\beta\mid \boldsymbol{y}\mid +\gamma\mid \boldsymbol{z}\mid}{r_1} \tag{26}$$

To satisfy condition (26), $\boldsymbol{s}$ is stable in the sense of Lyapunov. According to (26) s can be decreased by increasing r_1. The upper bound of the error signal is determined from (26) by:

$$\mid \boldsymbol{e}\mid < \frac{r_1\mid \boldsymbol{s}\mid -\beta\mid \boldsymbol{y}\mid -\gamma\mid \boldsymbol{z}\mid}{\alpha} \tag{27}$$

with enough large r_1, the error dynamic becomes stable.

Though we can stabilize the error dynamic by increasing r_1, we can not ignore $\boldsymbol{q}_1$ (It can be very small). The existence of $\boldsymbol{q}_1$ determine the existence of chattering phenomenon. How to prevent chattering is an important aspect for sliding mode control. We know that the value of $\boldsymbol{q}_1$ in Eq. (18) determine the magnitude of chattering of controller. A fuzzy control system is introduced here to weaken the chattering. The input is $s\dot{s}$, the product of $\boldsymbol{s}$ and it's derivative $\dot{\boldsymbol{s}}$. The output is $\boldsymbol{q}_1(t)$, which substitute the value of $\boldsymbol{q}_1$. There are two rules to follow: (1) If $\boldsymbol{s}\dot{\boldsymbol{s}} > 0$, then $\boldsymbol{q}_1(t)$ increase; (2) If $\boldsymbol{s}\dot{\boldsymbol{s}} < 0$, then $\boldsymbol{q}_1(t)$ decrease.

We can design a fuzzy system about $\boldsymbol{s}\dot{\boldsymbol{s}}$ and the change $\boldsymbol{\Delta q}_1(t)$ of $\boldsymbol{q}_1(t)$, where $\boldsymbol{s}\dot{\boldsymbol{s}} = (s\dot{s}1, s\dot{s}2, \ldots, s\dot{s}n)^T$, $\boldsymbol{\Delta q}_1(t) = diag(\Delta q_{11}(t), \ldots, \Delta q_{1n}(t))$. The universes of discourse of $s\dot{s}j$ and $\Delta q_{1j}(t)$ are $[-3k_1, 3k_1], [-3k_2, 3k_2]$, respectively, where k_1, k_2, k_3 are positive constants.

Each universe of discourse is decomposed into seven fuzzy partitions, which are NB, NM, NS, ZO, PS, PM, PB. The fuzzy rules are as follows:

Rule 1: If $s\dot{s}j$ is NB then $\Delta q_{1j}(t)$ is NB;
Rule 2: If $s\dot{s}j$ is NM then $\Delta q_{1j}(t)$ is NM;
Rule 3: If $s\dot{s}j$ is NS then $\Delta q_{1j}(t)$ is NS;
Rule 4: If $s\dot{s}j$ is ZO then $\Delta q_{1j}(t)$ is ZO;
Rule 5: If $s\dot{s}j$ is PS then $\Delta q_{1j}(t)$ is PS;
Rule 6: If $s\dot{s}j$ is PM then $\Delta q_{1j}(t)$ is PM;
Rule 7: If $s\dot{s}j$ is PB then $\Delta q_{1j}(t)$ is PB;

With singleton fuzzifier, product inference engine and a weighted average defuzzification, we establish the fuzzy system, which is described by

$$\Delta q_{1j}(t) = \frac{\sum_{i=1}^{7} \bar{y}^{ij} \mu_{s\dot{s}ij}}{\sum_{i=1}^{7} \mu_{s\dot{s}ij}} \quad (j = 1, \ldots, n) \tag{28}$$

where $\bar{y}^{ij}$ is the singleton control in the i*th*-rule, $\mu_{s\dot{s}i}$ is the membership degree of $s\dot{s}$.

We get the value of $q_{1j}(t)$ by integration method:

$$q_{1j}(t) = G \int_0^t \Delta q_{1j}(t)dt \tag{29}$$

where G is a positive constant to give.

Then Eq. (20) is changed into

$$\boldsymbol{w}_1(t) = -(\boldsymbol{CK})^{-1}(\boldsymbol{CB}_1\boldsymbol{e} + \boldsymbol{q}_1(t)sgn(\boldsymbol{s}) + r_1\boldsymbol{s}) \tag{30}$$

From the foregoing discussion, we can weaken the chattering by increasing the value of r_1 (meanwhile decreasing the value of $\boldsymbol{q}_1$) and by using fuzzy system output to describe the switch gain $\boldsymbol{q}_1$.

5 Numerical Results and Analysis

In what follows, we illustrate the active fuzzy sliding mode control approach for UFPS of chaotic systems. A fourth order Runge-Kutta solver with time step size of 0.001 s is performed to solve the differential equations. Let the parameters of universe about fuzzy system $k_1 = 5,\ k_2 = 0.5$. Let $G = 5$ in Eq. (29). The membership function of fuzzy input $s\dot{s}j$ is shown in Fig. 1. The membership function of fuzzy output $\Delta q_{1j}(t)$ is shown in Fig. 2. The fuzzy system is established by using fuzzy logic toolbox.

Here the UFPS of hyperchaotic Chen system (drive system) [27] and hyperchaotic Liu system introduced in [28] (response system) in the sense of hyperchaotic Chen-Lee system [29] (reference system) is considered by using the prescribed active fuzzy sliding mode control method. The drive, response and reference systems are respectively given by

$$\begin{cases} \dot{x}_1 = 35(x_2 - x_1) + x_4 \\ \dot{x}_2 = 7x_1 - x_1x_3 + 12x_2 \\ \dot{x}_3 = x_1x_2 - 3x_3 \\ \dot{x}_4 = x_2x_3 + 0.7x_4 \end{cases} \tag{31}$$

Fig. 1 The membership function of fuzzy input $s\dot{s}$

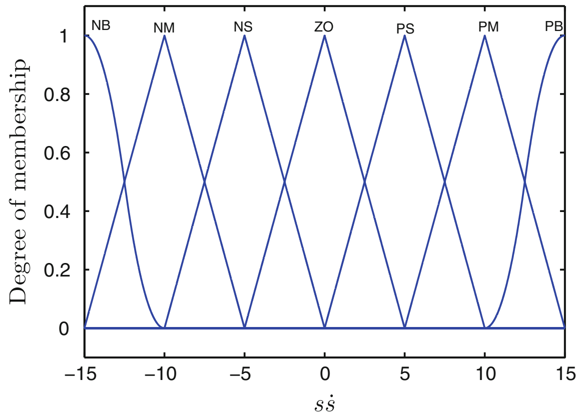
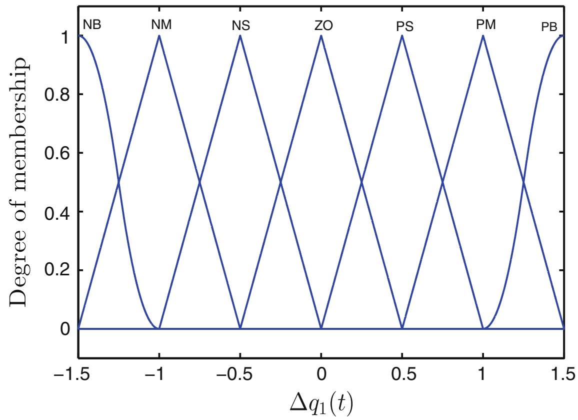


Fig. 2 The membership function of fuzzy output $\Delta q_1(t)$



$$\begin{cases} \dot{y}_1 = 10(y_2 - y_1) + u_1 \\ \dot{y}_2 = 40y_1 - y_1y_3 + y_4 + u_2 \\ \dot{y}_3 = -2.5y_3 + 4y_1^2 + u_3 \\ \dot{y}_4 = -20y_1 + u_4 \end{cases} \tag{32}$$

$$\begin{cases} \dot{z}_1 = -z_2z_3 + 5z_1 \\ \dot{z}_2 = z_1z_3 - 10z_2 \\ \dot{z}_3 = 1/3z_1z_2 - 3.8z_3 + 0.2z_4 \\ \dot{z}_4 = 1.2z_1 + 0.5z_2z_3 + 0.05z_4 \end{cases} \tag{33}$$

Initial conditions of the drive, response, and reference systems are $\boldsymbol{x}_0 = (1, 2, 3, 5)^T$, $\boldsymbol{y}_0 = (-10, 3, 8, 7)^T$ and $\boldsymbol{z}_0 = (2, 1, 6, 5)^T$, respectively. Let us assume

$\boldsymbol{A}(t) = diag(sin(t) + 2, 0.5sin(t) + 2, 0.5sin(t) + 4, cos(t) + 2)$. The control parameters are chosen as $r_1 = 200$. Let $\boldsymbol{K} = diag(-10, -30, -10, -20)$. Let constant matrix $\boldsymbol{C} = diag(4, 4, 4, 4)$ in the sliding surface Eq. (15). Let the total time $T = 20$ s.

Uncertainties will be considered in the three system dynamics in matrices of form:

$$\Delta \boldsymbol{B}_1 = diag(0.09, 0.06, 0.11, 0.07),$$
$$\Delta \boldsymbol{B}_2 = diag(-0.02, 0.05, 0.01, 0.03),$$
$$\Delta \boldsymbol{B}_3 = diag(0.05, 0.04, -0.02, 0.03).$$

Let us assume the uncertain nonlinear parts as

$$\Delta \boldsymbol{g}_1(\boldsymbol{x}) = (sin\, x_1, 0.5sin\, x_2, sin^2 x_1, 0.2sin\, x_2\, sin\, x_4)^T$$
$$\Delta \boldsymbol{g}_2(\boldsymbol{y}) = (sin\, y_2, 0.5sin\, y_1 cos\, y_2, 0.2sin\, y_3, 0.1sin^2 y_4)^T$$
$$\Delta \boldsymbol{g}_3(\boldsymbol{z}) = (0.4sin\, z_3, 0.1sin\, z_1 sin\, z_2, 0.3sin\, z_3, sin\, z_4)^T$$

The phase portrait of the drive and response system are shown in Figs. 3 and 4.

For $r_1 = 200$, the errors reach 0 in a very short time. For clarity, we only draw the time evolution diagram of the UFPS errors in 0.05 s which are depicted in Fig. 5.

Figure 5 displays $\boldsymbol{e} \to 0$ with $t \to \infty$. No more than 0.03 s, $\boldsymbol{e}$ reaches 0. We know UFPS achieved with our designed control law.

The time evolution of $\boldsymbol{u}$ are depicted in Fig. 6. From Fig. 6, we know the chattering phenomenon which is inherent to a sliding mode control is avoided by smoothing the switch signal via fuzzy system.

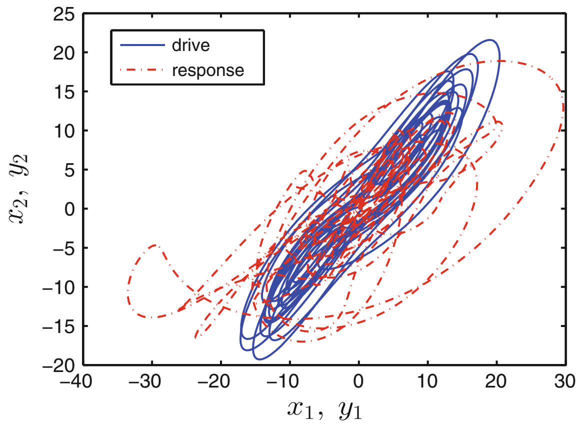


Fig. 3 The phase diagram of the drive and response systems on $x_1 - x_2(y_1 - y_2)$ plane

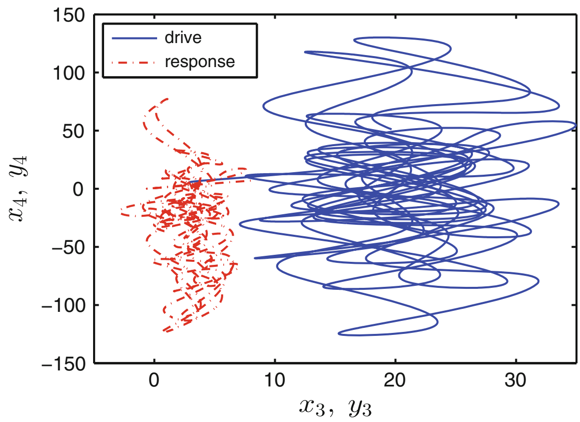


Fig. 4 The phase diagram of the drive and response systems on $x_3 - x_4(y_3 - y_4)$ plane

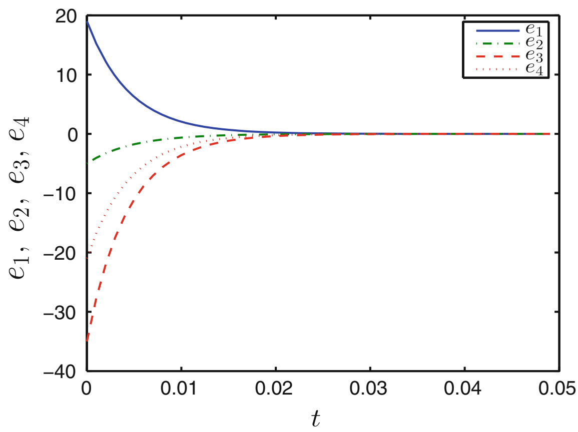


Fig. 5 The time evolution of the errors e_1, e_2, e_3, e_4

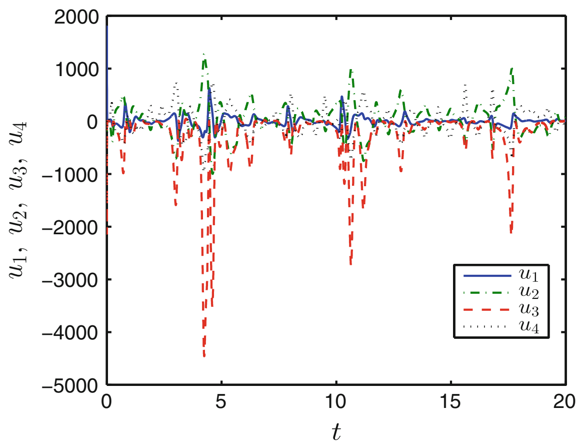


Fig. 6 The time evolution of the controller u_1, u_2, u_3, u_4

6 Conclusion

In this paper, we have recalled the definition UFPS. It involves three systems. An active fuzzy sliding mode control method has been given for the UFPS of chaotic systems. UFPS of different chaotic systems in the sense of chaotic system was successfully accomplished. The chaotic systems with uncertainty are considered. Simulation results verified the effectiveness of the proposed scheme. It was shown that the prescribed fuzzy system reduced the chattering dramatically.

Acknowledgments Thanks to the support by National Natural Science Foundation of China (No. 61374118) and Scientific Research Fund Project of Yunnan Provincial Department of Education (No. 2014Y441).

References

1. Pecora, L.M., Carroll, T.L.: Synchronization in chaotic systems. Phys. Rev. Lett. **64**(8), 821–824 (1990)
2. Ahn, C.K., et al.: Adaptive synchronization for uncertain chaotic systems with external disturbance. Commun. Nonlinear Sci. Numer. Simul. **15**(8), 2168–2177 (2010)
3. Rulkov, N.F., Sushchik, M.M., Tsimring, L.S., Abarbanel, H.D.: Generalized synchronization of chaos in directionally coupled chaotic systems. Phys. Rev. E **51**(2), 980–994 (1995)
4. Cao, L.Y., Lai, Y.C.: Antiphase synchronism in chaotic systems. Phys. Rev. E **58**(1), 382–386 (1998)
5. Ahn, C.K.: An approach to anti-synchronization for chaotic systems. Phys. Lett. A **373**(20), 1729–1733 (2009)
6. Ahn, C.K.: Adaptive H^{∞} anti-synchronization for time-delayed chaotic neural networks. Prog. Theoret. Phys. **122**(6), 1391–1403 (2009)
7. Mainieri, R., Rehacek, J.: Projective synchronization in three-dimensional chaotic systems. Phys. Rev. Lett. **82**(15), 3042–3045 (1999)
8. Jia, Q.: Projective synchronization of a new hyperchaotic Lorenz system. Phys. Lett. A **370**(1), 40–45 (2007)
9. Li, G.H.: Modified projective synchronization of chaotic system. Chaos, Solitons Fractals **32**(5), 1786–1790 (2007)
10. Wang, X., Zhang, X., Ma, C.: Modified projective synchronization of fractional-order chaotic systems via active sliding mode control. Nonlinear Dyn. **69**(1–2), 511–517 (2012)
11. Chen, Y., Li, X.: Function projective synchronization between two identical chaotic systems. Int. J. Mod. Phys. C **18**(05), 883–888 (2007)
12. Tang, X., Lu, J., Zhang, W.: The funciton projective synchronization of chaotic system using backstepping design. J. Dyn. Control **5**, 216–219 (2007)
13. Du, H., Zeng, Q., Wang, C.: Function projective synchronization of different chaotic systems with uncertain parameters. Phys. Lett. A **372**(33), 5402–5410 (2008)
14. Du, H., Zeng, Q., Wang, C.: Modified function projective synchronization of chaotic system. Chaos, Solitons Fractals **42**(4), 2399–2404 (2009)
15. Sebastian Sudheer, K., Sabir, M.: Switched modified function projective synchronization of hyperchaotic Qi system with uncertain parameters. Commun. Nonlinear Sci. Numer. Simul. **15**(12), 4058–4064 (2010)
16. Zheng, S.: Adaptive modified function projective synchronization of unknown chaotic systems with different order. Appl. Math. Comput. **218**(10), 5891–5899 (2012)

17. Zhang, B., Li, H.: Universal projective synchronization of two different hyperchaotic systems with unknown parameters. J. Appl. Math. **2014** (2014)
18. Utkin, V., Guldner, J., Shijun, M.: Sliding Mode Control in Electro-mechanical Systems, vol. 34. CRC Press (1999)
19. Zhang, H., Ma, X.K., Liu, W.Z.: Synchronization of chaotic systems with parametric uncertainty using active sliding mode control. Chaos, Solitons Fractals **21**(5), 1249–1257 (2004)
20. Haeri, M., Tavazoei, M.S., Naseh, M.R.: Synchronization of uncertain chaotic systems using active sliding mode control. Chaos, Solitons Fractals **33**(4), 1230–1239 (2007)
21. Tavazoei, M.S., Haeri, M.: Synchronization of chaotic fractional-order systems via active sliding mode controller. Phys. A: Stat. Mech. Appl. **387**(1), 57–70 (2008)
22. Fu-Cai, L., Jia-Qiu, S.: Anti-synchronizing different chaotic systems using active sliding mode control. Acta Phys. Sin. **57**(8), 4729–4737 (2008)
23. Zhao, Y.: Synchronization of two coupled systems of JJ type using active sliding mode control. Chaos, Solitons Fractals **42**(5), 3035–3041 (2009)
24. Cai, N., Jing, Y., Zhang, S.: Modified projective synchronization of chaotic systems with disturbances via active sliding mode control. Commun. Nonlinear Sci. Numer. Simul. **15**(6), 1613–1620 (2010)
25. Yahyazadeh, M., Ranjbar Noei, A., Ghaderi, R.: Synchronization of chaotic systems with known and unknown parameters using a modified active sliding mode control. ISA Trans. **50**(2), 262–267 (2011)
26. Wang, L.-X..: A course in fuzzy systems and control. Prentice-Hall International, Inc. (1999)
27. Li, Y., Tang, W.K., Chen, G.: Generating hyperchaos via state feedback control. Int. J. Bifurcat. Chaos **15**(10), 3367–3375 (2005)
28. Wang, F.-Q., Liu, C.-X.: Hyperchaos evolved from the Liu chaotic system. Chin. Phys. **15**(5), 963–968 (2006)
29. Chen, C.H., Sheu, L.J., Chen, H.K., Chen, J.H., Wang, H.C., Chao, Y.C., Lin, Y.K.: A new hyper-chaotic system and its synchronization. Nonlinear Anal.: Real World Appl. **10**(4), 2088–2096 (2009)

Hesitant Fuzzy Prioritized Hybrid Average Operator and Its Application to Multiple Attribute Decision Making

Chuan-yang Ruan, Jian-hui Yang, Li-na Han, Jing Duan and Ruo-bing Liu

Abstract Considering the double impact of priority level of attribute and the dispersion degree of hesitant fuzzy elements, we propose a prioritized hybrid weighted method based on the hesitant fuzzy information entropy. Then, we put forward the hesitant fuzzy prioritized hybrid average (HFPHA) operator based on the hybrid weighted methods, furthermore, we discuss some desirable properties of the proposed operator.

Keywords Hesitant fuzzy sets · Integration operator · Priority level · Information entropy

1 Introduction

Hesitant fuzzy set (HFS) is a generalized fuzzy set that is extended from fuzzy set (FS) by Torra [1]. The significant characteristic is to allow different evaluation values on one evaluation index simultaneously, HFS has been further applied to many fields at present [2–7]. A key decision technique of hesitant fuzzy set is hesitant fuzzy information aggregation, especially weighted operator aggregation with preference information. It's easy to find that a key factor in the hesitant fuzzy weighted operators is weight, so the calculation of the weight of aggregation operator has drawn attention of many scholars [8–17]. For instance, O'Hagan [12] proposed a maximum entropy weighted method. Ahn [13] presented a method when preference relationships between alternatives are specified. Fuller and Majlender [14] proposed a minimum variance

C. Ruan · J. Yang (✉) · J. Duan · R. Liu
School of Business Administration, South China University of Technology, Guangzhou 510641, China
e-mail: bmjhyang@scut.edu.cn

L. Han
School of Economics and Commerce, South China University of Technology, Guangzhou 510006, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_22

weighted method. Wang and Parkan [15] proposed a minimax disparity approach for obtaining aggregation operator weights. On the basis of intuitionistic fuzzy information, Zhou [16] proposed an accurate weight-determined (AWD) method relying on the membership and non-membership and Wang and Parkan [17] proposed a preemptive goal programming method for obtaining aggregation operator weights. We can see that current methods are under the assumption that the attributes are at the same priority level. However, in real and practical information aggregation problem, the attributes always have different priority level. Therefore, we investigate a new weight-determined method based on the hesitant fuzzy information: a prioritized hybrid weighted method based on the hesitant fuzzy information entropy. The core idea of the method is to calculate the entropy of attribute, and give greater weight to attributes that having less number of hesitant fuzzy elements and more average (i.e., the more unified expert opinions), in the meantime, maintains original priority relationships between different attribute values, which means unity and monotony. On the basis of the prioritized hybrid weighted method, we propose a hesitant fuzzy weighted aggregation operator: hesitant fuzzy prioritized hybrid average (HFPHA) operator, and provide the characteristics and operation rules of the new operators. Finally, validity and superiority of the operator is confirmed by numerical examples.

2 Fundamental Theory of Hesitant Fuzzy Set

Zadeh [18] first attempted to express the uncertainty of information by the membership function, one type of the mathematical methods, however, there exists a lot of uncertainty in many practical decision-making operators. In order to deal with the element that one evaluative index emerging different values, Torra [1] proposed the concept of hesitant fuzzy set.

Definition 1 [1, 19] If there exists a non-empty set $X = \{x_1, x_2, \ldots, x_n\}$, the hesitant fuzzy set is a function of a subset from $X = \{x_1, x_2, \ldots, x_n\}$ to [0,1], to be easily understood, which can be described by mathematical symbol as follows:

$$E = \{\langle x, h_E(x)\rangle; |x \in X\} \quad (1)$$

where $h_E(x)$ indicates a set of some possible values in [0,1], $x \in X$ indicates a set of membership degree in set E.

Definition 2 [19] If there exists a non-empty hesitant fuzzy sets h, $s(h) = \frac{1}{\#h}\sum_{\gamma \in h} \gamma$ is called the score function of h, where $\#h$ is the number of the elements in hesitant fuzzy set h. For two hesitant fuzzy sets h_1 and h_2, if $s(h_1) \succ s(h_2)$, then $h_1 \succ h_2$; if $s(h_1) = s(h_2)$, then $h_1 = h_2$.

Based on the characteristics of hesitant fuzzy set and the operational rules of intuitionistic fuzzy set, Xia [19] defined some new operational rules of the hesitant fuzzy sets h, h1 and h2.

(1) $h^{\lambda} = \cup_{\gamma \in h}\{\gamma^{\lambda}\}$
(2) $\lambda h = \cup_{\gamma \in h}\left\{1-(1-\gamma)^{\lambda}\right\}$
(3) $h_1 \oplus h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2}\{\gamma_1 + \gamma_2 - \gamma_1\gamma_2\}$
(4) $h_1 \otimes h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2}\{\gamma_1\gamma_2\}$

Considering the influence of priority level, on the basis of prioritized average (PA) [20], Wei [21] puts forward the hesitant fuzzy prioritized weighted average (HFPWA) and weighted geometric (HFPWG) operators, which are defined as follows:

Definition 3 [21] Given a collection of hesitant fuzzy elements $h_j (j = 1, 2, \ldots, n)$, then:

$$\begin{aligned} HFPWA(h_1, h_2, \ldots, h_n) &= \frac{T_1}{\sum_{j=1}^{n} T_j} h_1 \oplus \frac{T_2}{\sum_{j=1}^{n} T_j} h_2 \oplus \cdots \oplus \frac{T_n}{\sum_{j=1}^{n} T_j} h_n \\ &= \bigoplus_{j=1}^{n} \left(\frac{T_j h_j}{\sum_{j=1}^{n} T_j}\right) \end{aligned}$$

is defined as hesitant fuzzy prioritized weighted average (HFPWA) operator. where $T_j = \prod_{k=1}^{j-1} s(h_k) (j = 2, 3, \ldots, n), T_1 = 1, s(h_k)$ is the score values of h_k.

3 Hesitant Fuzzy Prioritized Weighted Method

Hesitant fuzzy information aggregation is one of the core contents of hesitant fuzzy set theory. However, most scholars did not consider the prioritized relationships between the assembled attributes, and assumed independent among the attributes. In the real world, there always exists a priority relationship among the attributes, and the elements number in hesitant fuzzy set will influence the credibility of evaluated data. Therefore, we first propose a hybrid weighted method based on the hesitant fuzzy information entropy considering the priority level of attributes, then construct a hesitant fuzzy information aggregation operator based on the hybrid weighted method: hesitant fuzzy prioritized hybrid average (HFPHA) operator.

3.1 Hesitant Fuzzy Information Entropy

The steps of calculating hesitant fuzzy information entropy are as follows [22, 23]:

(1) To quantify the hesitant fuzzy information within each attribute, calculate the weight w_{js} of the sth hesitant fuzzy number x_{js} of the jth attribute.

$$w_{js} = \frac{x_{js}}{\sum_{s=1}^{p} x_{js}} \quad (3)$$

(2) To calculate the entropy value e_j of the jth attribute.

$$e_j = -k \sum_{s=1}^{p} w_{js} \ln w_{js} \quad (4)$$

where k > 0, ln is the natural logarithm, $e_j \geq 0$. If values of the jth attribute are equal, then $w_{js} = \frac{x_{js}}{\sum_{s=1}^{p} x_{js}} = \frac{1}{p}$, and e_j is the biggest, that is $e_j = -k \sum_{s=1}^{p} \frac{1}{p} \ln \frac{1}{p} = k \ln p$, If $k = \frac{1}{\ln p}$, then $0 \leq e_j \leq 1$.

For the jth attribute, the less discrepancies of x_{js}, the bigger value of e_j; When values of the jth attribute are equal, the experts' opinions are completely unified, only one data is retained in the hesitant fuzzy set, then let $e_j = e_{\max} = 1$; The more value discrepancies of the jth attribute, the smaller value of e_j, then the lower credibility degree of the data in this attribute.

3.2 A Prioritized Hybrid Weighted Method Based on the Hesitant Fuzzy Information Entropy

We present a nonlinear hybrid weighted method, which not only ensures the priority level of the attribute, but also contains the data information. Detailed steps are as follows:

(1) The decision makers give preference information of the attribute, i.e. the attribute prioritization.
(2) Determine the proportion r_j of significance degree of attributes x_j and x_{j+1} at the adjacent priority level by calculating the hesitant fuzzy information entropy (see Sect. 3.1).

$$r_j = \begin{cases} e_j / e_{j+1}, \mathrm{e_j} \geq e_{j+1} (\mathrm{j} = 1, 2, \ldots, \mathrm{n} - 1) \\ 1, \mathrm{e_j} < \mathrm{e_{j+1}} (\mathrm{j} = 1, 2, \ldots, \mathrm{n} - 1) \end{cases} \quad (5)$$

(3) Based on the given r_j, calculate the weight t_n of the nth attribute in the lowest priority level:

$$t_n = \left(1 + \sum_{k=2}^{n} \prod_{i=k}^{n} r_i\right)^{-1} \tag{6}$$

(4) Calculate the weights in other priority levels through the following formula with the weight t_n:

$$t_{n-1} = r_n t_n \tag{7}$$

where t_n indicates the weight of the nth attribute in the lowest priority level.

4 Hesitant Fuzzy Prioritized Hybrid Average (HFPHA) Operator

Based on the existing hesitant fuzzy information aggregation operator and the prioritized hybrid weighted method based on the hesitant fuzzy information entropy, we propose the hesitant fuzzy prioritized hybrid average (HFPHA) operator.

Definition 4 Given a collection of hesitant fuzzy elements $h_j (j = 1, 2, \ldots, n)$, and let HFPHA: $\Omega_n \to \Omega$. If

$$HFPHA(h_1, h_2, \ldots, h_n) = t_1 h_1 \oplus t_2 h_2 \oplus \cdots \oplus t_n h_n \tag{8}$$

where $t_j (j = 1, 2 \ldots, n)$ is the prioritized hybrid weight based on the hesitant fuzzy information entropy of the jth attribute (see Sect. 3.1 in detail), and $t_j \in [0, 1]$, $\sum_{j=1}^{n} t_j = 1$, then HFPHA is called the hesitant fuzzy prioritized weighted average operator.

According to the algorithm of the hesitant fuzzy set, we can get the following theorems:

Theorem 1 *Given a collection of hesitant fuzzy elements* $h_j (j = 1, 2, \ldots, n)$, $t_j (j = 1, 2 \ldots, n)$ *is the prioritized hybrid weight based on the hesitant fuzzy information entropy of the jth attribute, their aggregated value by using the HFPHA operator is also a hesitant fuzzy element, and*

$$HFPHA(h_1, h_2, \ldots, h_n) = \cup_{\gamma_j \in h_j} (1 - \prod_{j=1}^{n} (1 - \gamma_j)^{t_j}) \tag{9}$$

Proof This theorem can be proved by mathematical induction

When $n = 2$, since $t_j h_j = \cup_{\gamma_j \in h_j}(1-(1-\gamma_j)^{t_j})$ we can obtain, $HFPHA(h_1, h_2) = t_1 h_1 \oplus t_2 h_2 = \cup_{\gamma_1 \in h_1, \gamma_2 \in h_2}\left\{(1-(1-\gamma_1)^{t_1}(1-\gamma_2)^{t_2}\right\}$ that is, when $n = 2$, the original formula holds.

If when $n = k$, the original formula holds, that is,

$$HFPHA(h_1, h_2, \ldots, h_k) = \cup_{\gamma_j \in h_j}(1 - \prod_{j=1}^{k}(1-\gamma_j)^{t_j})$$

then, when $n = k + 1$,

$$\begin{aligned} HFPHA(h_1, h_2, \ldots, h_{k+1}) &= \cup_{\gamma_j \in h_j}(1 - \prod_{j=1}^{k}(1-\gamma_j)^{t_j}) \oplus t_{k+1}h_{k+1} \\ &= \cup_{\gamma_j \in h_j}(1 - \prod_{j=1}^{k}(1-\gamma_j)^{t_j}) \oplus \cup_{\gamma_{k+1} \in h_{k+1}}(1-(1-\gamma_{k+1})^{t_{k+1}}) \\ &= \cup_{\gamma_j \in h_j}(1 - \prod_{j=1}^{k+1}(1-\gamma_j)^{t_j}) \end{aligned}$$

that is, when $n = k + 1$, the original formula holds too, then the original formula holds for any n.

5 Personnel Evaluation Based on the HFPHA Operator

Let us consider a real and practical problem. there are five sales engineers probation end dates are approaching, and now we must evaluate their performances to determine whether their contracts should be renewed. A strict evaluation of each of the five employees $A_j (j = 1, 2, 3, 4, 5)$ is performed from four aspects: work attitude G_1, communication skill G_2, problem solving skill G_3, and learning skill G_4. The prioritization of the criteria can be expressed as $G_1 \succ G_2 \succ G_3 \succ G_4$, where "$\succ$" indicates "priority to" [24]. Four decision makers $E = \{e_1, e_2, e_3, e_4\}$ evaluate the five alternatives $A_i (i = 1, 2, 3, 4, 5)$ under the above four indicators $G_j (j = 1, 2, 3, 4)$ anonymously, If two decision makers provide the same value, then delete the repeated values, so the value will emerge only once, and the evaluation values can be represented by the hesitant fuzzy set $h_{ij} (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$, the hesitant fuzzy decision matrix $H = (h_{ij})_{m \times n}$ is shown as follows (Table 1).

In order to obtain the ranking of alternatives, the detailed steps are as follows:

Step 1 Calculate the prioritized hybrid weights $t_{ij} (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$ based on the hesitant fuzzy information entropy, as indicated in Sect. 3.1.

Table 1 Hesitant fuzzy decision matrix

	G_1	G_2	G_3	G_4
A_1	{0.4, 0.5, 0.7}	{0.5, 0.8}	{0.6, 0.7, 0.9}	{0.5, 0.6}
A_2	{0.6, 0.7, 0.8}	{0.5, 0.6}	{0.4, 0.6, 0.7}	{0.4, 0.5}
A_3	{0.6, 0.8}	{0.2, 0.3,0.5}	{0.4, 0.6}	{0.5, 0.7}
A_4	{0.5, 0.6, 0.7}	{0.4, 0.5}	{0.8, 0.9}	{0.3, 0.4, 0.5}
A_5	{0.6, 0.7}	{0.5, 0.7}	{0.7, 0.8}	{0.2, 0.3, 0.4}

$$t_{ij} = \begin{bmatrix} 0.253 & 0.249 & 0.249 & 0.249 \\ 0.252 & 0.252 & 0.248 & 0.248 \\ 0.259 & 0.247 & 0.247 & 0.247 \\ 0.251 & 0.251 & 0.251 & 0.247 \\ 0.255 & 0.251 & 0.251 & 0.243 \end{bmatrix}$$

Step 2 Aggregate the ith row of the hesitant fuzzy matrix $H = (h_{ij})_{m \times n}$ by the HFPHA operator, obtain the integrated performance value $h_i (i = 1, 2, \ldots, 5)$ of the alternative A_i. As there are too many values in h_i, so we take the integrated performance value h_1 of the alternatives A_1 as an example, the other values are similar as h_1.

$$h_1 = \{0.5047, 0.5270, 0.5844, 0.6057, 0.6235, 0.6692, 0.5389, 0.5597, 0.6131, 0.6330, 0.6495, 0.6920, 0.6493, 0.6651, 0.7057, 0.7208, 0.7334, 0.7657, 0.5315, 0.5526, 0.6068, 0.6270, 0.6439, 0.6870, 0.5639, 0.5835, 0.6340, 0.6528, 0.6685, 0.7087, 0.6682, 0.6832, 0.7216, 0.7359, 0.7478, 0.7784\}$$

Step 3 Calculate the following scores of $h_i (i = 1, 2, \ldots, 5)$ according to the formula of the hesitant fuzzy function in Definition 2.

$s(h_1) = 0.6454, s(h_2) = 0.5840, s(h_3) = 0.5657, s(h_4) = 0.6321, s(h_5) = 0.6115$ then $s(h_1) \succ s(h_4) \succ s(h_5) \succ s(h_2) \succ s(h_3)$ the bigger the value of $s(h_i)$, the better the alternative A_i, thus the best alternative is A_1.

6 Conclusion

In order to differentiate and aggregate the information from different experts in group decision making in the hesitant fuzzy environment, we propose a prioritized hybrid weighted method based on the hesitant fuzzy information entropy and put forward HFPHA operator based on the prioritized hybrid weighted method, which considers the double impact of priority level of the attribute and the dispersion degree of hesitant fuzzy elements. The above characteristics could guarantee a constant priority level of attribute and distinguish the unity degree of expert opinions effectively, better reflect the inherent relationships among the hesitant fuzzy elements.

References

1. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
2. Xu, Z.S., Xia, M.M.: Distance and similarity measures for hesitant fuzzy sets. Inform. Sci. **181** (11), 2128–2138 (2011)
3. Mesut, Y., Basar, O., Cevik, O.S., et al.: Multi-criteria evaluation of alternative-fuel vehicles via a hierarchical hesitant fuzzy linguistic model. Expert Syst. Appl. **42**(5), 2835–2848 (2015)
4. Beg, I., Rashid, T.: TOPSIS for hesitant fuzzy linguistic term sets. Int. J. Intell. Syst. **28**, 1162–1171 (2013)
5. Rodriguez, R.M., Martinez, L., Torra, V., et al.: Hesitant Fuzzy Sets: State of the Art and Future Directions. Int. J. Intell. Syst. **29**, 495–524 (2014)
6. Farhadinia, B.: A series of score functions for hesitant fuzzy sets. Inf. Sci. **277**, 102–110 (2014)
7. Wang, J.Q., Wu, J.T.: Method for multi-criteria decision-making with hesitant fuzzy linguistic based on outranking relation. Control Decis. **30**(5), 887–891 (2015)
8. Ahn B.S: Parameterized OWA operator weights: an extreme point approach. Int. J. Approximate Reasoning **51**, 820–831 (2010)
9. Emrouznejad, A.: MP-OWA: the most preferred OWA operator. Knowl.-Based Syst. **21**, 847–851 (2008)
10. Liu, X.: On the properties of equidifferent OWA operator. Int. J. Approximate Reasoning **43**, 90–107 (2006)
11. Majlender, P.: OWA operators with maximal Rényi entropy. Fuzzy Sets Syst. **155**, 340–360 (2005)
12. O'Hagan, M.: Aggregating template rule antecedents in real time expert systems with fuzzy set logic. In: Proceedings of the 22nd Annual IEEE Asilomar Confon Signals, Systems and Computers. California, pp. 681–689 (1988)
13. Ahn, B.S.: Preference relation approach for obtaining OWA operators weights. Int. J. Approximate Reasoning **47**, 166–178 (2008)
14. Fuller, R., Majlender, P.: On obtaining minimal variability OWA operator weights. Fuzzy Sets Syst. **136**(2), 203–215 (2003)
15. Wang, Y.M., Parkan, C.: A minimax disparity approach for obtaining OWA operator weights. Inform. Sci. **175**(1/2), 20–29 (2005)
16. Zhou, W., He, J.M., Yu, D.J.: Accurate method of obtaining decision expert weights in intuitionistic fuzzy group decision making. Control Decis. **28**(5), 716–725 (2013)
17. Wang, Y.M., Parkan, C.: A preemptive goal programming method for aggregating OWA operator weights in decision making. Inform. Sci. **177**(8), 1867–1877 (2007)
18. Zadeh, L.A.: Fuzzy sets. Inform. Control **8**(3), 338–353 (1965)
19. Xia, M.M., Xu, Z.S.: Hesitant fuzzy information aggregation in decision making. Int. J. Approximate Reasoning **52**, 395–407 (2011)
20. Yager, R.R.: Prioritized aggregation operators. Int. J. Approximate Reasoning **48**, 263–274 (2008)
21. Wei, G.W.: Hesitant fuzzy prioritized operators and their application to multiple attribute decision making. Knowl.-Based Syst. **31**, 176–182 (2012)
22. Li, G.: The science and technology evalution model and its empirical research based on entropy_revised G1 combination weighting. Soft Sci. **24**(5), 31–36 (2010)
23. Ruan, C.Y., Yang, J.H.: Software quality evaluation model based on weighted mutation rate correction incompletion G1 combination weights. Math. Prob. Eng. **2014**, Article ID 541292 (2014)
24. Yu, D.J., Zhang, W.Y., Xu, Y.J.: Group decision making under hesitant fuzzy environment with application to personnel evaluation. Knowl.-Based Syst. **52**, 1–10 (2013)

Differential Transform Method for Solving Linear System of First-Order Fuzzy Differential Equations

Lei Wang and Na Li

Abstract In this paper we study an approximate-analytical method for the linear systems of first-order fuzzy differential equations (FDEs) with fuzzy initial value conditions under Seikkala derivative. The differential transform method (DTM) is used for the approximate-analytical solution of this problem. This method is illustrated by numerical examples.

Keywords Differential transform method · System of first-order linear FDEs · Approximate-analytical solution

1 Introduction

Fuzzy differential equations (FDEs) have been suggested as a way of modelling uncertain and incompletely specified systems and were studied by many researchers (see [1–3]). The concept of fuzzy derivative was first introduced by Chang and Zadeh [4]. It was followed up by Dubois and Prade [5], who defined and used the extension principle. A comprehensive approach to FDEs has been the work of Seikkala [6], and Kaleva [7], especially in its generalized form given by Buckley and Feuring [8]. For further investigation, the fuzzy dynamical systems based on FDEs are also widely applied to fuzzy control systems [9], bifurcations of fuzzy nonlinear dynamical systems [10], chaotic systems [11, 12], engineering [13] and many other fields (see [14, 15]). Similar to ordinary differential equations (ODEs), the exact analytical solutions of FDEs are often difficult, thus approximate–analytical methods for solving FDEs are of particular importance.

In this paper, an approximate-analytical method to the linear systems of first-order fuzzy differential equations with fuzzy initial conditions is presented. The

L. Wang (✉) · N. Li
Department of Basic Teaching, Liaoning Technical University,
Huludao 125105, China
e-mail: wllyq78@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_23

structure of the paper is organized as follows: In Sect. 2, some basic definitions and notations which will be used are brought. In Sect. 3, the proposed method DTM for solving the linear systems of first-order FDEs is presented. In Sect. 4, the proposed method DTM is illustrated by numerical examples. Conclusion is drawn in Sect. 5.

2 Preliminaries

We now recall some definitions needed through the paper. The basic definition of fuzzy numbers is given in [5].

By R, we denote the set of all real numbers. A fuzzy number is a mapping $u\colon R \to [0, 1]$ with the following properties:

(a) u is upper semi-continuous,
(b) u is fuzzy convex, i.e., $u(\lambda x + (1-\lambda)y) \geq min\{u(x), u(y)\}$ for all $x, y \in R, \lambda \in [0, 1]$,
(c) u is normal, i.e., $\exists x_0 \in R$ for which $u(x_0) = 1$,
(d) supp $u = \{x \in R | u(x) > 0\}$ is the support of the u, and its closure cl(supp u) is compact.

Let E be the set of all fuzzy number on R. The r-level set of a fuzzy number $u \in E, o \leq r \leq 1$, denoted by $[u]_r$, is defined as

$$[u]_r = \begin{cases} \{x \in R | u(x) \geq r\} & if \quad 0 \leq r \leq 1, \\ cl(supp\, u) & if \quad r = 0. \end{cases}$$

It is clear that the r-level set of a fuzzy number is a closed and bounded interval $[\underline{u}(r), \overline{u}(r)]$, where $\underline{u}(r)$ denotes the left-hand endpoint of $[u]_r$ and $\overline{u}(r)$ denotes the right-hand endpoint of $[u]_r$.

Definition 1 [16] A fuzzy number u in parametric form is a pair $(\underline{u}, \overline{u})$ of functions $\underline{u}(r), \overline{u}(r),\ 0 \leq r \leq 1$, which satisfy the following requirements:

- $\underline{u}(r)$ is a bounded non-decreasing left continuous function in (0.1], and right continuous at 0,
- $\overline{u}(r)$ is a bounded non-increasing left continuous function in (0.1], and right continuous at 0,
- $\underline{u}(r) \leq \overline{u}(r),\ 0 \leq r \leq 1.$

Definition 2 [6] Let $u = (\underline{u}(r), \overline{u}(r)), v = (\underline{v}(r), \overline{v}(r))$, then the Hausdorff distance between fuzzy numbers given by

$$D\colon E \times E \to R_+ \bigcup 0,$$

$$D(u, v) = \sup_{r \in [0,1]} max\{|\underline{u}(r) - \underline{v}(r)|, |\overline{u}(r) - \overline{v}(r)|\}.$$

Lemma 1 [17] *Let* $u = (\underline{u}(r), \overline{u}(r)), v = (\underline{v}(r), \overline{v}(r))$, *then*

$$u + v = (\underline{u}(r) + \underline{v}(r), \overline{u}(r) + \overline{v}(r)),$$
$$u - v = (\underline{u}(r) - \overline{v}(r), \overline{u}(r) - \underline{v}(r)),$$
$$k \times u = \begin{cases} (k\underline{u}, k\overline{u}), k \geq o, \\ (k\overline{u}, k\underline{u}), k < o. \end{cases}$$

Definition 3 [6] Let I be a real interval. A mapping $v: I \to E$ is called a fuzzy process and we denoted the r-level set by $[v(t)]_r = [\underline{v}(t,r), \overline{v}(t,r)]$. The Seikkala derivative $v^{'}(t)$ of $v(t)$ is defined by

$$[v^{'}(t)]_r = [\underline{v}^{'}(t,r), \overline{v}^{'}(t,r)]$$

provided that is a equation defines a fuzzy number $v(t) \in E$.

Recently, a method has been presented to transform an FDE to a system of ODEs [18]. Based on this method, a nonhomogeneous n-dimensional system of first-order linear FDEs is defined by

$$x^{'}(t) = Ax(t) + f(t), \quad x(0) = x_0 \tag{1}$$

where $A = (a_{ij}) \in R^{n \times n}, x = (x_1, x_2, \ldots x_n), x_0 = (x_{10}, x_{20}, \ldots x_{n0}) \in E^n, t \in T \subset R$ and $f = (f_1, f_2, \ldots, f_n): T \to E^n$. According to Definition 3, the system (1) is transformed into a system of ODEs. Let $[x(t)]_r = [\underline{x}(t,r), \overline{x}(t,r)]$. If $x(t)$ is Seikkala differentiable, then $[x^{'}(t)]_r = [\underline{x^{'}}(t,r), \overline{x^{'}}(t,r)]$. Thus, (1) is transformed into the following $2n \times 2n$ system of ODEs:

$$\begin{cases} \begin{pmatrix} \underline{x}(t,r) \\ \overline{x^{'}}(t,r) \end{pmatrix} = P \begin{pmatrix} \underline{x}(t,r) \\ \overline{x}(t,r) \end{pmatrix} + \begin{pmatrix} \underline{f}(t,r) \\ \overline{f}(t,r) \end{pmatrix}, \\ \begin{pmatrix} \underline{x}(0,r) \\ \overline{x}(0,r) \end{pmatrix} = \begin{pmatrix} \underline{x_0}(r) \\ \overline{x_0}(r) \end{pmatrix}. \end{cases} \tag{2}$$

where $P = (p_{ij}) \in R^{2n \times 2n}, p_{ij}$ are determined as follows:

$$\begin{cases} p_{ij} = a_{ij}, & a_{ij} \geq 0, \\ p_{i+n,j+n} = a_{ij}, & a_{ij} \geq 0, \\ p_{i,j+n} = a_{ij}, & a_{ij} < 0, \\ p_{i+n,j} = a_{ij}, & a_{ij} < 0, \\ p_{ij} = 0, & otherwise. \end{cases} \tag{3}$$

3 Differential Transform Method

The concept of differential transformation was introduced first by Zhou [19], and it was applied to solve linear and nonlinear initial value problems in electric circuit analysis. With this technique, it is possible to obtain highly accurate results or exact solutions for differential equations (see [20–22]).

Definition 4 If $x(t,r)$ is Seikkala differentiable of order k in time domain T then

$$\underline{\varphi}(t,k,r) = \frac{d^k(\underline{x}(t,r))}{dt^k} \quad \forall t \in T,$$
$$\underline{X}_i(k,r) = \underline{\varphi}(t_i,k,r) = \frac{d^k(\underline{x}(t,r))}{dt^k}\Big]_{t=t_i} \quad \forall k \in K$$
$$\overline{\varphi}(t,k,r) = \frac{d^k(\overline{x}(t,r))}{dt^k} \quad \forall t \in T,$$
$$\overline{X}_i(k,r) = \overline{\varphi}(t_i,k,r) = \frac{d^k(\overline{x}(t,r))}{dt^k}\Big]_{t=t_i} \quad \forall k \in K.$$

where k belongs to the set of non-negative integer, denotes as the K, $\underline{X}(k,r)$ and $\overline{X}(k,r)$ are called the lower and the upper spectrum of $X(t,r)$ at $t = t_i$ in the domain K, respectively.

Using the differential transform, Eq. (2) can be described as

$$\begin{cases} \begin{pmatrix} \underline{X}(k+1,r) \\ \overline{X}(k+1,r) \end{pmatrix} = \frac{1}{k+1} P \begin{pmatrix} \underline{X}(k,r) \\ \overline{X}(k,r) \end{pmatrix} + \begin{pmatrix} \underline{F}(k,r) \\ \overline{F}(k,r) \end{pmatrix}, \\ \begin{pmatrix} \underline{X}(0,r) \\ \overline{X}(0,r) \end{pmatrix} = \begin{pmatrix} \underline{x_0}(r) \\ \overline{x_0}(r) \end{pmatrix}. \end{cases} \tag{4}$$

So, $x(t,r)$ can be represented as

$$\underline{x}(t,r) = \sum_{k=0}^{\infty} \frac{(t-t_i)^k}{k!} \underline{X}(k,r),$$
$$\overline{x}(t,r) = \sum_{k=0}^{\infty} \frac{(t-t_i)^k}{k!} \overline{X}(k,r).$$

It can be easily proven that the transformation function has basic mathematical operations shown in Table 1.

Table 1 Basic mathematical properties of differential transform

Functional form	Differential transform
$z(x) = u(x) \pm v(x)$	$W(k) = U(k,h) \pm V(k,h)$
$z(x) = \alpha u(x)$	$W(k,h) = \alpha U(k,h)$
$z(x) = \left(\frac{d^m u(x)}{dx^m}\right)$	$W(k) = \frac{(m+k)!}{k!} U(k+m)$
$z(x) = u(x) \cdot v(x)$	$W(k) = \sum_{i=0}^{k} U(k-i)V(i)$
$w(x) = x^m$	$\delta(k-m)$
$w(x) = \exp(\lambda x)$	$W(k) = \lambda^k / k!$

4 Example

Example 1 Consider the following system of first-order linear FDEs (see [23])

$$\frac{d}{dt}\begin{pmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \end{pmatrix} + \begin{pmatrix} \tilde{1} \\ \tilde{2} \end{pmatrix} \tag{5}$$

Let us assume that $\tilde{x}_1(0) = \tilde{x}_2(0)$ to be a fuzzy number about 100, here $[\tilde{1}]_r = [r, 2-r]$, $[\tilde{2}]_r = [1+r, 3-r]$ and $[\tilde{x}_1(0)]_r = [\tilde{x}_2(0)]_r = [70+30r, 130-30r]$, where $r \in [0,1]$.

From Eqs. (4) and (5), we have

$$\frac{d}{dt}\tilde{x}(t,r) = P\tilde{x}(t,r) + \tilde{f}(t,r), \tag{6}$$

Here, $\tilde{x}(t,r) = \begin{pmatrix} \underline{x}_1(t,r) \\ \underline{x}_2(t,r) \\ \overline{x}_1(t,r) \\ \overline{x}_2(t,r) \end{pmatrix}, P = \begin{pmatrix} 0 & 2 & -3 & 0 \\ 3 & 0 & 0 & -4 \\ -3 & 0 & 0 & 2 \\ 0 & -4 & 3 & 0 \end{pmatrix}$, and

$f(t,r) = \begin{pmatrix} r \\ 1+r \\ 2-r \\ 3-r \end{pmatrix}$, subject to the initial condition $\tilde{x}(0,r) = \begin{pmatrix} 70+30r \\ 70+30r \\ 130-30r \\ 130-30r \end{pmatrix}$ for all $r \in [0,1]$.

The exact solutions of problem (5) for $r = 0$ are as follows:

$$\begin{pmatrix} \underline{x}_1(t,0) \\ \underline{x}_2(t,0) \\ \overline{x}_1(t,0) \\ \overline{x}_2(t,0) \end{pmatrix} = \begin{pmatrix} -\frac{31}{5}e^{t} + \frac{1}{15}e^{-6t} - \frac{362}{15}e^{6t} + \frac{493}{5}e^{-t} + \frac{5}{3} \\ \frac{31}{5}e^{t} - \frac{1}{10}e^{-6t} - \frac{181}{5}e^{6t} + \frac{493}{5}e^{-t} + \frac{3}{2} \\ \frac{31}{5}e^{t} + \frac{1}{15}e^{-6t} + \frac{362}{15}e^{6t} + \frac{493}{5}e^{-t} + 1 \\ -\frac{31}{5}e^{t} - \frac{1}{10}e^{-6t} + \frac{181}{5}e^{6t} + \frac{493}{5}e^{-t} + \frac{3}{2} \end{pmatrix} \tag{7}$$

Table 2 Numerical results for $\underline{x}_2(t, 0)$

t	Exact solution	Approximate solution	Relative error
0.1	70	70	0
0.2	40.0943635	41.5543635	4.787×10^{-3}
0.3	–5.284587	–5.284587	4.9296×10^{-2}
0.4	–79.6449203	–80.6449203	2.0116×10^{-1}
0.5	–207.509032	–207.609032	4.504×10^{-1}
0.6	–433.4791208	–434.2791208	5.18810^{-1}
0.7	–838.7532122	–838.7532122	1.18002

Table 3 Numerical results for $\bar{x}_2(t, 0)$

t	Exact solution	Approximate solution	Relative error
0.1	130	130	0
0.2	149.770729	149.7714901	7.6×10^{-4}
0.3	194.812268	194.809755	2.513×10^{-3}
0.4	285.15626	285.164812	8.55×10^{-3}
0.5	457.374157	457.307083	6.7074×10^{-2}
0.6	778.173309	778.085417	8.7892×10^{-2}
0.7	1369.169045	1369.365584	1.9654×10^{-1}

Using the proposed method DTM, the following approximate solutions are obtained:

$$\begin{pmatrix} \underline{x}_1(t,0) \\ \underline{x}_2(t,0) \\ \underline{x}_1(t,0) \\ \underline{x}_2(t,0) \end{pmatrix} = \begin{pmatrix} 70 - 309t - \frac{1201}{2}t^2 - \frac{3950}{3}t^3 - \frac{23489}{12}t^4 - \frac{70321}{30}t^5 \\ 70 - \frac{309}{2}t - 200t^2 - \frac{659}{2}t^3 - \frac{3919}{10}t^4 - \frac{46979}{120}t^5 \\ 130 + 52t + 489t^2 + \frac{2558}{3}t^3 + \frac{7877}{6}t^4 + \frac{23428}{15}t^5 \\ 130 + 113t + \frac{1395}{2}t^2 + \frac{3872}{3}t^3 + \frac{23483}{12}t^4 + \frac{70627}{30}t^5 \end{pmatrix}. \tag{8}$$

The above-mentioned approximate solutions are, in fact, fifth order polynomial approximations of the exact solutions. Numerical and graphical comparisons between the exact solutions and the approximate solutions for $r = 0$ are given in Tables 2 and 3, and Figs. 1 and 2.

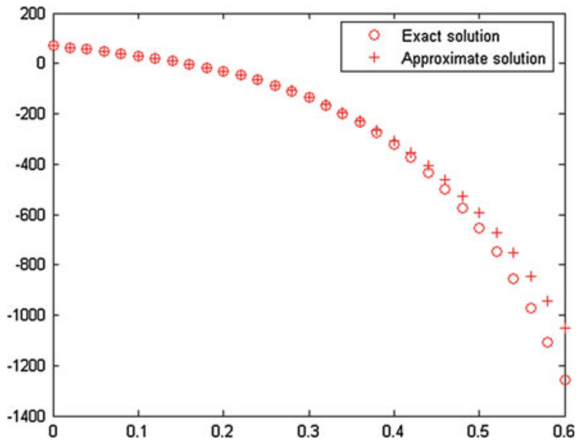


Fig. 1 The results for $\underline{x}_2(t, 0)$

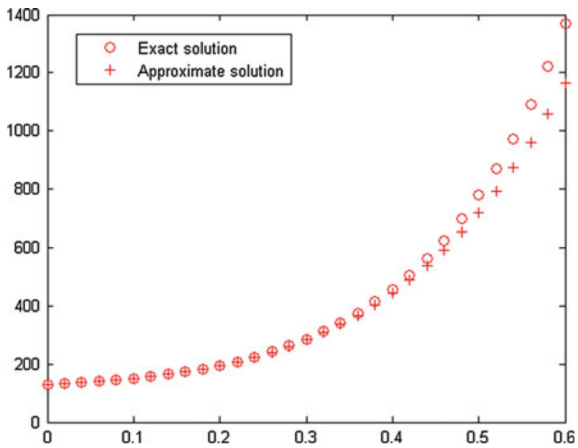


Fig. 2 The results for $\bar{x}_2(t, 0)$

5 Conclusion

In this paper the differential transform method(DTM) for solving the linear systems of first-order fuzzy differential equations with fuzzy initial conditions is presented. The numerical results obtained by present method are compared with the analytical solutions.

Acknowledgments This work was supported by a grant from the Ph.D. Programs Foundation of Ministry of Education of China (No. 20102121110002), the National Natural Science Foundations of China (Grants No. 61304173), Foundation of Liaoning Educational Committee (Grant No. 13-1069). The authors are most grateful to the referees for valuable suggestions and comments.

References

1. Nieto, J.J., Rodríguez-López, R., Franco, D.: Linear first-order fuzzy differential equations. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **14**, 687–709 (2006)
2. Bede, B., Rudas, I.J., Bencsik, A.L.: First order linear fuzzy differential equations under generalized differentiability. Inform. Sci. **177**, 1648–1662 (2007)
3. Allahviranloo, T., Ahmady, N., Ahmady, E.: Numerical solution of fuzzy differential equations by predictor-corrector method. Inform. Sci. **177**, 1633–1647 (2007)
4. Chang, S.L., Zadeh, L.A.: On fuzzy mapping and control. IEEE Trans. Syst. Man Cybern. **2**, 30–34 (1972)
5. Dubois, D., Prade, H.: Toward fuzzy differential calculus: Part 3, differentiation. Fuzzy Sets Syst. **8**, 225–233 (1982)
6. Seikkala, S.: On the fuzzy initial value problem. Fuzzy Sets Syst. **24**, 319–330 (1987)
7. Kaleva, O.: The Cauchy problem for fuzzy differential equations. Fuzzy Sets Syst. **35**, 389–396 (1990)
8. Buckley, J.J., Feuring, T.: Fuzzy differential equations. Fuzzy Sets Syst. **110**, 43–54 (2000)
9. Chen, B., Liu, X.: Reliable control design of fuzzy dynamical systems with time-varying delay. Fuzzy Sets Syst. **146**, 349–374 (2004)
10. Hong, L., Sun, J.Q.: Bifurcations of fuzzy nonlinear dynamical systems. Commun. Nonlinear Sci. Numer. Simul. **11**, 1–12 (2006)
11. Román-Flores, H., Chalco-Cano, Y.: Robinsons chaos in set-valued discrete systems. Chaos Solitons Fractals **25**, 33–42 (2005)
12. Román-Flores, H., Chalco-Cano, Y.: Some chaotic properties of Zadehs extension. Chaos Solitons Fractals **35**, 112–119 (2008)
13. Song, S.J., Guo, L., Feng, C.: Global existence of solutions to fuzzy differential equations. Fuzzy Sets Syst. **115**, 371–376 (2000)
14. Xu, J.P., Liao, Z.G., Hu, Z.N.: A class of linear differential dynamical systems with fuzzy initial condition. Fuzzy Sets Syst. **157**, 2339–2358 (2007)
15. Xu, J.P., Liao, Z.G., Nieto, J.J.: A class of linear differential dynamical systems with fuzzy matrices. J. Math. Anal. Appl. **368**, 54–68 (2010)
16. Ma, M., Friedman, M., Kandel, A.: Numerical solution of fuzzy differential equations. Fuzzy Sets Syst. **105**, 133–138 (1999)
17. Ralescu, D.: A survey of the representation of fuzzy concepts and its applications, Advances in Fuzzy Set Theory and Applications, pp. 77–91. North-Holland, Amsterdam (1979)
18. Bede, B.: Note on "Numerical solutions of fuzzy differential equations by predictor–corrector method". Inform. Sci. **178**, 1917–1922 (2008)
19. Zhou, J.K.: Differential Transformation and its Applications for Electrical Circuits. Huazhong University press, (1986)
20. Erturk, V.S.: Differential transformation method for solving fuzzy differential equations of lane-emden type. Math. Comput. Appl. **12**, 135–139 (2007)
21. Allahviranloo, T., Kiani, N.A., Motamedi, N.: Solving fuzzy differential equations differential transformation method. Inform. Sci. **179**, 956–966 (2009)
22. Hassan, I.H.A.-H.: Application to differential transformation method for solving systems of differential equations. Appl. Math. Model. **32**, 2552–2559 (2008)
23. Buckley, J.J., Feuring, T., Hayashi, Y.: Linear system of first order ordinary differential equations: fuzzy initial equation. Soft. Comput. **6**, 415–421 (2002)

Weight of Basic Health Service Equalization Index Based on the Intuitionistic Fuzzy Analytic Hierarchy Process

Wei-xia Li , Zhuan-ku Huang, Zhong-ming Feng and Cheng-yi Zhang

Abstract According to the new consistency adjustment method of intuitionistic fuzzy complementary judgment matrix which is contained in the intuitionistic fuzzy analytic hierarchy process, intuitionistic fuzzy analytic hierarchy process was given. Then apply intuitionistic fuzzy analytic hierarchy process to get the weight of the index for the evaluation of basic health service equalization.

Keywords Intuitionistic fuzzy analytic hierarchy process · Intuitionistic fuzzy complementary judgment matrix · Consistency recursive iterative adjustment algorithm · Basic health service equalization · Weight

1 Introduction

As the generalization of fuzzy sets, IFS [1] is more practical, scientific and reasonable than fuzzy value in solving the fuzziness and the uncertainty problems [2–5], but it makes calculation more complicated and the operation process of this kind of matrix is very complex and difficult. So after the judgment matrix under the intuitionistic fuzzy environment was introduced [6, 7], there are a few research of intuitionistic fuzzy complementary judgment matrix (IFCJM), moreover there are aslo a few paper about the intuitionistic fuzzy analytic hierarchy process (IFAHP). This study firstly puts forward a method of consistency recursive iteration adjustment algorithm for the judgment matrix under the intuitionistic fuzzy environment. Then based on the consistency recursive iteration adjustment algorithm of IFCJM, this study introduces the process of IFAHP. Based on the current status of the increasing

W. Li
Department of School of Public Health,
Hainan Medical University, Haikou 571199, China

Z. Huang · Z. Feng · C. Zhang (✉)
Department of Mathematics and Statistics,
Hainan Normal University, Haikou 571158, China
e-mail: chengyizh@hainnu.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_24

Chinese residents requirement for health quality, the health service level should be studied, moreover, the health service equalization which is an important part of the health quality must be evaluated. This study firstly used the IFAHP method to get the weight which is used for the evaluation of basic health service equalization [8–10].

2 Consistency Adjustment Algorithm of Intuitionistic Fuzzy Complementary Judgment Matrix

2.1 IFCJM and Its Properties

Definition 1 Let $A = (a_{ij})_{n*n}$ be a judgement matrix, where $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})(i, j \in N)$, if $t_{ij} \in [0, 1]$, $f_{ij} \in [0, 1]$, $t_{ji} = f_{ij}, \pi_{ji} = \pi_{ij}, t_{ii} = f_{ii} = 0.5, t_{ij} + f_{ij} \leq 1$, then A is called IFCJM.

Note The meaning of the elements $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$ is as follows: t_{ij} represents importance degree of x_i relative to x_j, f_{ij} represents importance degree of x_j relative to x_i and π_{ij} represents the uncertainty importance degree of x_i relative to x_j. The scale of the elements t_{ij} and f_{ij} takes for 0.1–0.9 nine scales.

Definition 2 Let $A = (a_{ij})_{n*n}$ be an IFCJM where $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$, if for each k, $a_{ij} = a_{ik} - a_{jk}$, such that $t_{ij} = t_{ik} - t_{jk} + 0.5, f_{ij} = f_{ik} - f_{jk} + 0.5$ and $\pi_{ij} = \pi_{ik} - \pi_{jk}$, then is called ACIFCJM.

Definition 3 Let $A = (a_{ij})_{n*n}$ be an IFCJM, a_{kj} and a_{ij} be the elements of A, then $a_{ij} - a_{kj} = (t_{ij} - t_{kj} + 0.5, f_{ij} - f_{kj} + 0.5, \pi_{ij} - \pi_{kj})$ is called the subduction and $a_{ij} + a_{kj} = (t_{ij} + t_{kj} - 0.5, f_{ij} + f_{kj} - 0.5, \pi_{ij} + \pi_{kj})$ is called addition.

Definition 4 Let $b_{ij} = (tb_{ij}, fb_{ij})$ be intuitionistic fuzzy value, then $a_{ij} = f(b_{ij})$ is called the conversion equation from b_{ij} to $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$ such that

$$\begin{cases} t_{ij} = 0.1 + (tb_{ij} + 0.3)/3 \\ f_{ij} = 0.1 + (fb_{ij} + 0.3)/3 \\ \pi_{ij} = 1 - t_{ij} - f_{ij} \end{cases}$$

Definition 5 Let $B = (b_{ij})_{n*n}$ be ACIFCJM where $b_{ij} = (tb_{ij}, fb_{ij})$. If $tb_{ij} < 0$ or $fb_{ij} < 0$ or $tb_{ij} > 1$ or $fb_{ij} > 1$, then all the elements of B are adjusted for $p_{ij} = (tp_{ij}, fp_{ij}, \pi p_{ij})$ of $P = (p_{ij})_{n*n}$ such that $\begin{cases} tp_{ij} = 0.1 + (tb_{ij} + 0.3)/3 \\ fp_{ij} = 0.1 + (fb_{ij} + 0.3)/3 \\ \pi p_{ij} = 1 - t_{ij} - f_{ij} \end{cases}$. Then P is called the scale transition matrix of B.

2.2 The Consistency Adjustment Algorithm

2.2.1 Basic Definition and Theorem

Symbols are as follows:

1. Let $A = (a_{ij})_{n*n}$ be an IFCJM where $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$, then $A^{(k)}$ signifies the leading principal submatrix of order K of A.
2. Let $A^{(k)}$ be the leading principal submatrix of order K of A, then $A_k^{(s)}$ signifies the leading principal submatrix of order s of $A^{(k)}$ where $1 \le s \le k$.
3. Let $A^{(k)}$ be the leading principal submatrix of order K of A, then B^k signifies the ACIFCJM of $A^{(k)}$.
4. Let $A = (a_{ij})_{n*n}$ be an IFCJM, then $C_k^{(k-1)} = (c_{ij}^{(k-1)})_{k*k}$ signifies the leading principal submatrix of order K, which satisfied the leading principal submatrix of order k-1 is ACIFCJM $B^{(k-1)}$ and the elements in the kth row (column) are the same as A.

Definition 6 Let $A = (a_{ij})_{n*n}$ be an IFCJM, $B^{(k-1)}$ be the ACIFCJM of $A^{(k-1)}$. let $C_k^{(k-1)} = [tc_{ij}^{(k-1)}, fc_{ij}^{(k-1)}]_{k*k}$, if $I_k = \{m \mid |tc_{tm}^{(k-1)} - t_{km}| < 0.5, |fc_{tm}^{(k-1)} - f_{km}| < 0.5$, and $-0.9 \le (tc_{tm}^{(k-1)} - t_{km}) + (fc_{tm}^{(k-1)} - f_{km}) \le 0, \forall t \epsilon \{1, 2, \ldots, k\}$, then I_k is called the consistency index of the consistency adjustment of $A^{(k)}$ according to the elements in the kth row.

Definition 7 Let $R = ([tr_{ij}, fr_{ij}])_{k*k}$ and $P = ([tp_{ij}, fp_{ij}])_{k*k}$ be consistent IFCJM, then $e = (\sum_{i=1}^{k} \sum_{j=1}^{k} (|tr_{ij} - tp_{ij}| + |fr_{ij} - fp_{ij}|)/k$ is called deviate value of R and P.

2.2.2 The Additive Consistent Adjustment Algorithm of IFCJM

Let $A = (a_{ij})_{n*n}$ be IFCJM where $a_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$, then the additive consistent adjustment algorithm is as follows:

Step 1: $A_1 = (1)$ and $\begin{pmatrix} 1 & a_{12} \\ a_{21} & 2 \end{pmatrix}$ that are ACIFCJM are the leading principal submatrix of order one and order two of A respectively.

Step 2: Suppose for each $k > 2, h = 1, 2, \ldots, k-1$, all the $A_k^{(h)}$ have been adjusted for ACIFCJM. Then we adjust the elements in the kth row of $C_k^{(k-1)} = [tc_{ij}^{(k-1)}, fc_{ij}^{(k-1)}]_{k*k}$ whose leading principal submatrix of order $k-1$ equals $B^{(k)}$. If $I_i \ne \phi (i = 1, 2, \ldots, k-1)$, then $tb_{kt}^{(km)} = t_{km} - tc_{mt}^{(k-1)} + 0.5$, $fb_{kt}^{(km)} = f_{km} - fc_{mt}^{(k-1)} + 0.5$, $\pi b_{kt}^{(km)} = \pi_{km} - \pi c_{mt}^{(k-1)}$ and $tb_{tk}^{(km)} = fb_{kt}^{(km)}$, $fb_{tk}^{(km)} = tb_{kt}^{(km)}$ where $t \in \{1, 2, \ldots k\}$. Continue to step 3. Otherwise, if $\exists 1 \le h \le k-1$, such that $I_h = \phi$, then go to step 7.

Step 3: If $I_k \neq \phi$, then let $T_k^m = (t_{ij}^{km})_{k*k}$ where $m \in I_k$ and calculate deviate value e_k^m of $A^{(k)}$ and T_k^m. Continue to step 4. Otherwise, go to step 5.

Step 4: Determine $J_k = \{l_k \mid l_k \in I_k\}$ such that $e_k^{l_k} = \min\{e_k^m\}$, then let $B^{(k)} = T_k^{l_k} \mid l_k = \min_{l_k \in J_k}(l_k)$ and go to step 8.

Step 5: Let $T_k^m = (b_{ij}^{(km)})_{k*k}$ where $m \in \{1, 2, \ldots k\}$, then calculate deviate value e_k^m of $A^{(k)}$ and T_k^m.

Step 6: Determine $J_k = \{l_k \mid l_k \in I_k\}$ such that $e_k^{l_k} = \min\{e_k^m\}$, then get the scale transition matrix $S_k^{l_k}$ of $T_k^{l_k}$. Let $B^{(k)} = \{S_k^{l_k} \mid l_k = \min_{l_k \in J_k}(l_k)\}$, then go to step 8.

Step 7: If $\exists 1 \leq h \leq k-1$ such that $I_h = \phi$, then get $b_{km} = f(a_{km})$, $tb_{kt}^{(km)} = ta_{km} - tc_{mt}^{(k-1)} + 0.5$, $fb_{kt}^{(km)} = fa_{km} - fc_{mt}^{(k-1)} + 0.5$, and $tb_{tk}^{(km)} = fb_{kt}^{(km)}$, $fb_{tk}^{(km)} = tb_{kt}^{(km)}$ by the scale transition formula. Calculate the scale transition matrix $S^{(k)}$ of $A^{(k)}$, and let $A^{(k)} = S^{(k)}$, then go to step 3.

Step 8: Let $k = k + 1$; If $k \leq n$, then go to step 2. Otherwise, continue to step 9.

Step 9: let $B = B^{(k)}$, then output B.

Step 10: End.

2.2.3 Priority Vectors Formula of IFCJM

Definition 8 let $B = (b_{ij})_{n*n}$ be consistent IFCJM where $b_{ij} = (t_{ij}, f_{ij}, \pi_{ij})$, and $e = (t_e, f_e, \pi_e)$ represents the equal importance degree of x_i relative to x_j, then $b = (b_i)_{n*1}$ is called sum and normalized vector of line where $b_i = (t_i, f_i)$, $t_i = (\sum_{j=1}^{n}(t_{ij} - t_e))/n + t_e$, $f_i = (\sum_{j=1}^{n}(f_{ij} - f_e))/n + f_e$, $i = 1, 2, L, n$.

Theorem 1 *Let* $x = \langle u_A(x), v_A(x) \rangle$ *be intuitionistic fuzzy number of IFS A, then* $E_\lambda(x) = u_A(x) - v_A(x) + (2\lambda - 1)\pi_A(x)$ *which is based on the risk preference coefficient is sorting function of the intuitionistic fuzzy number, where* λ *is decision makers risk preference coefficient and* $\lambda \in [0, 1]$.

Definition 9 let $B = (b_{ij})_{n*n}$ be consistent IFCJM, $b = (b_i)_{n*1}$ be sum and normalized vector of line of B where $b_i = (t_i, f_i)$. Then $c = (c_i)_{n*1}$ which is based on the risk preference coefficient is called sorting of B intuitionistic fuzzy number of where $c_i = E_\lambda(b_i) = u_{b_i} - v_{b_i} + (2\lambda - 1)\pi_{b_i}$.

Definition 10 lets $B = (b_{ij})_{n*n}$ be consistent IFCJM, $c = (c_i)_{n*1}$ be sorting of intuitionistic fuzzy number of $B = (b_{ij})_{n*n}$, then $W = (w_i)_{n*1}$ is weight ordering vector of B and satisfies as follows:

(1) If $\exists c_i \leq 0$, then determine $R = (r_i)_{n*1}$ such that $r_i = 0.8c_i/(\max_i(c_i) - \min_i(c_i)) + (0.1 - 0.8\min_i(c_i))/(\max_i(c_i) - \min_i(c_i))$ and $w_i = r_i/\sum_{i=1}^{n} r_i$.

(2) If $\forall c_i > 0$, then $w_i = c_i/\sum_{i=1}^{n} c_i$.

3 Analytic Hierarchy Process Under the Intuitionistic Fuzzy Environment and Its Application

3.1 Intuitionistic Fuzzy Analytic Hierarchy Process

Suppose $V = (v_{ij})_{n*n}$ as IFCJM, it is usually given by decision makers personal experience and practical conditions and other factors. However, it is worthy of discussion how to avoid the subjective arbitrary of decision-makers to confirm the value of v_{ij}. In order to ensure v_{ij} more reasonable, decision-makers can ask k experienced experts to compare the attributes each other in the same level. Suppose $v_{ij}^q = (t_{ij}^q, f_{ij}^q)$ is given by the expert q, we can get the arithmetic mean of all the t_{ij}^q and all the f_{ij}^q respectively to summarize the opinions of experts, that is $t_{ij} = (\sum_{q=1}^{k} t_{ij}^q)/k$, $f_{ij} = (\sum_{q=1}^{k} f_{ij}^q)/k$, then $v_{ij} = (t_{ij}, f_{ij})$ can be confirmed, where $i, j = 1, 2, \ldots n$.

Step 1: By analyzing the relationship of different factors in the system, a systematic hierarchical structure can be established.

Step 2: By asking the experts to establish the IFCJM $V = (v_{ij})_{n*n}$ which is the pairwise comparisons of all the elements on the same level with respect to the element on the above level, where $v_{ij} = (t_{ij}, f_{ij})$, $i, j = 1, 2, \ldots n$, and t_{ij}, f_{ij} are given quantity standard on the scale of 0.1 to 0.9.

Step 3: Do the consistency adjustment for the IFCJM V by additive consistency recursive iterative adjustment algorithm. Then obtain the ACIFCJM.

Step 4: Calculate the priority vector. Calculate the weights of the elements on every level, and then calculate the combined weight of all the elements on the bottom level with respect to the element on the top level.

3.2 The Evaluation of Basic Health Service Equalization

Recently, people pay more attention on the evaluation of basic health service equalization, then a comprehensive evaluation model of basic health service equalization is constructed by using IFAHP.

Target layer	Primary index	Secondary index
The equalization of basic health service (O)	Health financing fairness ($A1$)	$B1$: Fiscal subsidies income about institution of disease prevention and control
		$B2$: Fiscal subsidies income about Health supervision institutions
		$B3$: Fiscal subsidies income about maternal and child health care
		$B4$: Surrounding greening
		$B5$: Basic medical insurance premiums of Urban
	Health resources allocation fairness($A2$)	$B6$: Technical personnel numbers of public health institutions
		$B7$: Public health institutions numbers
		$B8$: Beds numbers of public health institutions
	Health service fairness ($A3$)	$B9$: Water-improving rate in rural
		$B10$: The penetration of harmless toilets in the countryside
		$B11$: Care management rate of children under 7 years old
		$B12$: Rate of postpartum visit
	Health outcomes fairness ($A4$)	$B13$: Mortality rate of child under 5 years old
		$B14$: Maternal mortality rate
		$B15$: The incidence of low birth weight
		$B16$: Incidence of infectious diseases

Fig. 1 The evaluation index system

Table 1 IFCJM about the target larger

O	A1	A2	A3	A4
A1	(0.5, 0.5)	(0.2, 0.7)	(0.3, 0.6)	(0.6, 0.3)
A2	(0.7, 0.2)	(0.5, 0.5)	(0.6, 0.3)	(0.8, 0.1)
A3	(0.6, 0.3)	(0.3, 0.6)	(0.5, 0.5)	(0.7, 0.2)
A4	(0.3, 0.6)	(0.1, 0.8)	(0.2, 0.7)	(0.5, 0.5)

Step 1: Construct an evaluation index system and establish a hierarchical structure in Fig. 1.

Step 2: We ask the senior experts to give the evaluation values and get the IFCJM which is the pair-wise comparisons for all the primary indexes (or secondary indexes) with respect to the objective in the target layer (or primary indexes) as follows:

Table 1 IFCJM O which is the pair-wise comparison for all the primary indexes with respect to the objective in the target layer

Table 2 IFCJM which is the pair-wise comparison for all the secondary indexes with respect to in the primary indexes layer.

Similarly, we get the IFCJM $A2$, $A3$, $A4$.

Step 3: Do the consistency adjustment for the IFCJM, and get the weight vector.

Table 2 IFCJM about the primary indexes layer

A1	B1	B2	B3	B4	B5
B1	(0.5, 0.5)	(0.2, 0.7)	(0.1, 0.8)	(0.3, 0.6)	(0.1, 0.9)
B2	(0.7, 0.2)	(0.5, 0.5)	(0.3, 0.6)	(0.6, 0.4)	(0.1, 0.8)
B3	(0.8, 0.1)	(0.6, 0.3)	(0.5, 0.5)	(0.6, 0.3)	(0.2, 0.7)
B4	(0.6, 0.3)	(0.4, 0.6)	(0.3, 0.6)	(0.5, 0.5)	(0.3, 0.6)
B5	(0.9, 0.1)	(0.8, 0.1)	(0.7, 0.2)	(0.6, 0.3)	(0.5, 0.5)

$O^{(1)}$	$\begin{pmatrix} (0.3667,0.3667) & (0.2667,0.4333) & (0.3,0.4) & (0.4,0.3) \\ (0.4333,0.2667) & (0.3667,0.3667) & (0.4,0.333) & (0.5,0.2333) \\ (0.4,0.3) & (0.3333,0.4) & (0.3667,0.3667) & (0.4667,0.2667) \\ (0.3,0.4) & (0.2333,0.5) & (0.2667,0.4667) & (0.3667,0.3667) \end{pmatrix}$
$A^{(1)}$	$\begin{pmatrix} (0.3667,0.3667) & (0.2667,0.4333) & (0.2333,0.4667) & (0.3,0.4) & (0.1667,0.5667) \\ (0.4333,0.2667) & (0.3667,0.3667) & (0.3333,0.4) & (0.4,0.3333) & (0.2333,0.4667) \\ (0.4667,0.2333) & (0.4,0.3333) & (0.3667,0.3667) & (0.4333,0.3) & (0.2667,0.4333) \\ (0.4,0.3) & (0.3333,0.4) & (0.3,0.4333) & (0.3667,0.3667) & (0.2,0.5) \\ (0.5667,0.1667) & (0.4667,0.2333) & (0.4333,0.2667) & (0.5,0.2) & (0.3667,0.3667) \end{pmatrix}$

Fig. 2 ACLFCJM from IFCJM

Table 3 The weight sorting vector of ACIFCJM

α	0.5
$W^{(1)}$	$(0.1905, 0.4286, 0.3333, 0.0476)'$
$\omega^{(1)}$	$(0.0435, 0.1884, 0.2464, 0.1304, 0.3913)'$
$\omega^{(2)}$	$(0.0746, 0.2537, 0.6716)'$
$\omega^{(3)}$	$(0.1643, 0.45, 0.3357, 0.05)'$
$\omega^{(4)}$	$(0.45, 0.3, 0.2, 0.05)'$

(1) Do the consistency adjustment for the IFCJM, and get the ACIFCJM as follows in Fig. 2.
Similarly, we get the ACIFCJM $A^{(2)}, A^{(3)}, A^{(4)}$.
(2) Let decision makers risk preference coefficient $\alpha = 0.5$,then we get the weight sorting vector of the ACIFCJM in the following Table 3.

Step 4: Aggregate the global weight.

Let $\alpha = 0.5$, then we can get a weight matrix $W^{(1)}$ of all the secondary indexes on the third level with respect to all the primary indexes on the second level where $W^{(1)} = (\omega^{(1)}, \omega^{(2)}, \omega^{(3)}, \omega^{(4)}, \omega^{(5)})$. Aggregate the global weight. The combined weights of the elements in the secondary index level relative to the objective in the target layer is $W^{(2)} = W^{(1)} \cdot w^T = (0.0083, 0.0359, 0.0469, 0.0248, 0.0745, 0.032, 0.1087, 0.2878, 0.0548, 0.15, 0.1119, 0.0167, 0.0214, 0.0143, 0.0095, 0.0024)$.

4 Conclusion

Based on the consistency adjustment algorithm of IFCJM, introduced the IFAHP. According to the result of the IFAHP used in getting the weight of the health service index, we can do the evalution of the basic health service equalization next study in the future.

Acknowledgments This research is partly funded by the National Natural Science Foundation of China (NSFC) under Grant No. 71140008 & 71361008,and Special Application Technology Research and Development of Hainan province under Grant No. ZDXM2014107.

References

1. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
2. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
3. Zang, J.J.: Fuzzy analytic hierarchy process (FAHP). Fuzzy Syst. Math. **14**, 80–88 (2000)
4. Leung, L.C., Cao, D.: On consistency and ranking of alternatives in fuzzy AHP. Eur. J. Oper. Res **124**(1), 102C113 (2000)
5. Wang, Y.M., Fan, Z.P., Hua, Z.: A chi-square method for obtaining a priority vector from multiplicative and fuzzy preference relations. Eur. J. Oper. Res. **182**(1), 356C366 (2007)
6. Chen, J.M., Li, H.X.: A ranking method and decision-making with consistent risk preference in intuitionistic fuzzy numbers. Fuzzy Syst. Math. **24**(6), 85–91 (2010)
7. Wang, L.: The analytic hierarchy process based on vague sets and the application of the process, vol. 10. Hainan Normal University Master's Thesis (2011)
8. Wang, Z.H., Dong, Y.J., Bo, T.T., Cao, J.L., Fan, L.: Construction and quantitative analysis of basic service equalization index system. Chin. Health Econ. **33**(5), 45–48 (2014)
9. Mao, Y., Liu, J.J., Yang, J., Xu, F., Zhang, M.J.: Analysis on fairness of chinas human resource allocation for health in 2011. Chin. Health Econ. **32**(8), 35–38 (2013)
10. Wang, F., Zhu, X.L., Ding, X., Xu, F., Zhang, M.J.: Analysis the status of grassroots health human resource allocation in China and its equity. Chin. Health Serv. Manage. **2**, 108–110 (2012)

About Approach to Multi-attribute Decision Making Problems Based on COWA Operator Under Interval-Valued Intuitionistic Fuzzy Environment

Xiang-jun Xie and Xiao-xia Lv

Abstract Proposed new Hamming distance and entropy for interval-valued intuitionistic fuzzy sets. And then the two new measures, which are extended from the Hamming distance and entropy for intuitionistic fuzzy sets by using the continuous ordered weighted aggregation (COWA) operator, are applied in interval-valued intuitionistic fuzzy multi-attribute decision making (MADM) problems with attribute weights completely unknown. An example is illustrated to verify the effectiveness of the two measures. In fact, we can convert a MADM problem from interval-valued intuitionistic fuzzy environment to intuitionistic fuzzy environment first and then solve it by using the intuitionistic fuzzy MADM methods, which is a new strategy.

Keywords Interval-valued intuitionistic fuzzy sets · COWA operator · Hamming distance · Entropy · Multi-attribute decision making

1 Introduction

Considering the complexity of decision-making environment and the fuzziness of decision makers' thinking, Atanassov and Gargov extended the intuitionistic fuzzy sets [1] (IFSs) and proposed the theory of interval-valued intuitionistic fuzzy sets [2] (IVIFSs) in 1989. The IVIFSs, which describe the degree of membership and non-membership through interval numbers, has attracted a lot of attention since it's proposed. Much work has been done on its theories and applications. Xu and Chen [3–5] defined the arithmetic averaging operator, the weighted arithmetic averaging operator, the geometric averaging operator, the weighted geometric averaging

X. Xie (✉) · X. Lv
School of Science, Southwest Petroleum University,
Chengdu 610500, China
e-mail: xiangjunxie@126.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_25

operator, the ordered weighted geometric averaging operator and the hybrid geometric averaging operator and all operators are used to aggregate the information of an IVIFS. Ye [6] proposed a novel accuracy function and a new intuitionistic fuzzy MADM method. Li [7, 8] proposed the TOPSIS-based linear and non-linear programming method for MADM with IVIFSs. Li [9] proposed the generalized hybrid weighted aggregation operator and the continuous weighted aggregation operator for IVIFSs. The above-mentioned papers concentrate on MADM of IVIFSs with weights known or partially known. Work on cases with weights completely unknown is few. Ye [10], Gao and Wei [11] proposed MADM methods based on interval-valued intuitionistic fuzzy entropy. The attributes' weights are determined by entropy in both papers. We will study the MADM methodology using continuous ordered weighted aggregation (COWA) operator [12].

The rest of the paper are organized as follows. Section 2 introduce some preliminaries. In Sect. 3, the new Hamming distance and entropy for IVIFSs are proposed. Section 4 introduce work about MADM of IVIFSs. Section 5 is the conclusion.

2 Preliminaries

Definition 1 [2] An IVIFS $\tilde{A}$ in the set discourse X is given by:

$$\tilde{A} = \{\langle x, \tilde{u}_A(x), \tilde{v}_A(x)\rangle | x \in X\} = \left\{\langle x, \left[u_A^-(x), u_A^+(x)\right], \left[v_A^-(x), v_A^+(x)\right]\rangle | x \in X\right\}$$

where $\tilde{u}_A(x), \tilde{v}_A(x) \subseteq [0, 1]$ denote, respectively, the membership degree and non-membership degree of x to $\tilde{A}$, with the condition $u_A^+(x) + v_A^+(x) \le 1$ for $\forall x \in A$.

$\tilde{\pi}_A(x) = \left[\pi_A^-(x), \pi_A^+(x)\right] = \left[1 - u_A^+(x) - v_A^+(x), 1 - u_A^-(x) - v_A^-(x)\right]$ denotes the hesitancy of x to $\tilde{A}$.

Definition 2 [2] $\forall \tilde{A}, \tilde{B} \in IVIFS(X)$, the following operations are established as:
(1) $\tilde{A} \subseteq \tilde{B} \Leftrightarrow u_A^-(x) \le u_B^-(x), u_A^+(x) \le u_B^+(x)$ and $v_A^-(x) \ge v_B^-(x), v_A^+(x) \ge v_B^+(x)$;
(2) $\tilde{A} = \tilde{B} \Leftrightarrow \tilde{A} \subseteq \tilde{B}, \tilde{A} \supseteq \tilde{B}$;
(3) $\tilde{A}^C = \left\{\langle x, \left[v_A^-(x), v_A^+(x)\right], \left[u_A^-(x), u_A^+(x)\right]\rangle | x \in X\right\}$.

Theorem 1 [9] *$\tilde{A}, \tilde{B}, \tilde{C}$ are IVIFSs in discourse X, the distance between $\tilde{A}$ and $\tilde{B}$ satisfies the following conditions:*
(1) $0 \le d(\tilde{A}, \tilde{B}) \le 1$; (2) $d(\tilde{A}, \tilde{B}) = 0 \Leftrightarrow \tilde{A} = \tilde{B}$;
(3) $d(\tilde{A}, \tilde{B}) = d(\tilde{B}, \tilde{A})$; (4) $d(\tilde{A}, \tilde{B}) \le d(\tilde{A}, \tilde{C}) + d(\tilde{C}, \tilde{B})$.

3 New Hamming Distance and Entropy Based on COWA Operator for IVIFSs

Definition 3 [12] The mapping $f^c : \Gamma \to R$ is called as the continuous ordered weighted aggregation (COWA) operator, if: $f^c([a,b]) = \int_0^1 \frac{d\rho(y)}{dy}[b - y(b-a)]dy$, $[a,b] \in \Gamma$, where $\rho(y)$ is a basic-unit monotonic (BUM) function (i.e. the function satisfies conditions: $\rho(0) = 0$; $\rho(1) = 1$; $\rho(x) \geq \rho(y)$ if $x > y$ for $\forall x, y \in [0,1]$).

Le $\lambda = \int_0^1 \rho(y)dy$ (called as attitudinal character), then $f^c[a,b] = \lambda b + (1-\lambda)a$ (i.e. the weighted average of the interval endpoints value). So we can determine the result value by choosing different BUM function.

It's obvious that an IVIFS $\tilde{A} = \left\{\langle x, \left[u_A^-(x), u_A^+(x)\right], \left[v_A^-(x), v_A^+(x)\right]\rangle | x \in X\right\}$ will be aggregated into an IFS $A = \{\langle x, u_A(x), v_A(x)\rangle | x \in X\}$ (written as: $f^c(\tilde{A}) = A$), where $\pi_A(x) = \lambda \pi_A^-(x) + (1-\lambda)\pi_A^+(x)$ (written as $f_1^c[a,b] = \lambda a + (1-\lambda)b$).

Theorem 2 *Let* $f^c(\tilde{A}) = A$, $f^c(\tilde{B}) = B$, *the following conditions are satisfied:*

(1) $\tilde{A} = \tilde{B} \Rightarrow A = B$; (2) $\tilde{A} = \tilde{B}^C \Rightarrow A = B^C$; (3) $\tilde{A} \underline{\prec} \tilde{B} \Rightarrow A \underline{\prec} B$; (4) $\tilde{A} \subseteq \tilde{B} \Rightarrow A \subseteq B$.

3.1 Generalized Hamming Distance for IVIFS

Szmidt and Kacprzyk [13] pointed out that the distance measure between two IFSs should take three parameters (i.e. the membership degree, non-membership degree and hesitancy). They defined the Hamming distance between two IFSs as follows: the normalized Hamming distance:

$$d(A,B) = \frac{1}{2n}\sum_{i=1}^{n} |u_A(x_i) - u_B(x_i)| + |v_A(x_i) - v_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|$$

the weighted Hamming distance:

$$d(A,B)_w = \frac{1}{2n}\sum_{i=1}^{n} w_i[|u_A(x_i) - u_B(x_i)| + |v_A(x_i) - v_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|] \quad (1)$$

Based on the COWA operator, we will extend the Hamming distance of IFSs.

Definition 4 Let $\tilde{A}$ and $\tilde{B}$ are two IVIFSs in discourse $X = \{x_1, x_2, \ldots, x_n\}$, the Hamming distance measures are defined as follows:
the normalized Hamming distance:

$$d(\tilde{A},\tilde{B}) = \frac{1}{2n}\sum_{i=1}^{n} \left| f^c\left[u_A^-(x_i), u_A^+(x_i)\right] - f^c\left[u_B^-(x_i), u_B^+(x_i)\right]\right| + \left| f^c\left[v_A^-(x_i), v_A^+(x_i)\right] - f^c\left[v_B^-(x_i), v_B^+(x_i)\right]\right| + \left| f_1^c\left[\pi_A^-(x_i), \pi_A^+(x_i)\right] - f_1^c\left[\pi_B^-(x_i), \pi_B^+(x_i)\right]\right|$$

the weighted Hamming distance:

$$d(\tilde{A},\tilde{B})_w = \frac{1}{2n}\sum_{i=1}^{n} w_i\Big[\Big|f^c\big[u_A^-(x_i),u_A^+(x_i)\big] - f^c\big[u_B^-(x_i),u_B^+(x_i)\big]\Big| + \Big|f^c\big[v_A^-(x_i),v_A^+(x_i)\big] - f^c\big[v_B^-(x_i),v_B^+(x_i)\big]\Big| + \Big|f_1^c\big[\pi_A^-(x_i),\pi_A^+(x_i)\big] - f_1^c\big[\pi_B^-(x_i),\pi_B^+(x_i)\big]\Big|\Big] \tag{2}$$

we just need to verify the formulas satisfy the characters in Theorem 1. Le $f^c(\tilde{A}) = A$, $f^c(\tilde{B}) = B$, based on the Theorem 2, the Hamming distance between $\tilde{A}$ and $\tilde{B}$ can be viewed as Hamming distance between A and B. So the characters in Theorem 1 will be satisfied.

3.2 Generalized Entropy for IVIFS

Entropy measures the indeterminacy of information and it should take two aspects (i.e. fuzziness, lack of knowledge or non-specificity) [14] into consideration while handling the practical MADM problems. Gao and Sun [15] defined the following entropy formula:

$$E(A) = \frac{1}{n}\sum_{i=1}^{n}\frac{1-|u_A(x_i)-v_A(x_i)|^2+\pi_A^2(x_i)}{2} \tag{3}$$

Similarly, the entropy for IFSs will be extended.

Definition 5 Let $\tilde{A}$ be an IVIFS in $X = \{x_1, x_2, \ldots, x_n\}$, the entropy for $\tilde{A}$ is:

$$E(\tilde{A}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1-\big|f^c\big[u_A^-(x_i),u_A^+(x_i)\big]-f^c\big[v_A^-(x_i),v_A^+(x_i)\big]\big|^2+\big(f_1^c\big[\pi_A^-(x_i),\pi_A^+(x_i)\big]\big)^2}{2} \tag{4}$$

Theorem 3 *$E(\tilde{A})$ in formula (4) have the following properties:*
(1) $E(\tilde{A}) = 0 \Leftrightarrow \tilde{A}$ *is a crisp set;* (2) $E(\tilde{A}) = 0 \Leftrightarrow \big[\pi_A^-(x), \pi_B^-(x)\big] = [1,1]$;
(3) $E(\tilde{A}) = E(\tilde{A}^C)$; (4) $\tilde{A} \preceq \tilde{B} \Rightarrow E(\tilde{A}) \le E(\tilde{B})$.

Proof (1) $\tilde{A}$ is a crisp set, so $\tilde{u}_A(x_i) = [0,0]$, $\tilde{v}_A(x_i) = [1,1]$ or $\tilde{u}_A(x_i) = [1,1]$, $\tilde{v}_A(x_i) = [0,0]$, then $E(\tilde{A}) = 0$; if $E(\tilde{A}) = 0$, since $\big(f_1^c\big[\pi_A^-(x_i),\pi_A^+(x_i)\big]\big)^2 \ge 0$ and $\big|f^c\big[u_A^-(x_i),u_A^+(x_i)\big]-f^c\big[v_A^-(x_i),v_A^+(x_i)\big]\big|^2 \le 1$, so $\big(f_1^c\big[\pi_A^-(x_i),\pi_A^+(x_i)\big]\big)^2 = 0$ and $\big|f^c\big[u_A^-(x_i),u_A^+(x_i)\big]-f^c\big[v_A^-(x_i),v_A^+(x_i)\big]\big|^2 = 1$, namely $\tilde{u}_A(x_i) = [0,0]$,

$\tilde{v}_A(x_i) = [1, 1]$ or $\tilde{u}_A(x_i) = [1, 1]$, $\tilde{v}_A(x_i) = [0, 0]$, so $E(\tilde{A}) = 0$;
(2) and (3) are obviously established;
(4) if $f^c(\tilde{A}) = A, f^c(\tilde{B}) = B$, then $\tilde{A} \preceq \tilde{B} \Rightarrow A \preceq B$, so

$$E(\tilde{A}) = E(A) \le E(B) = E(\tilde{B}).$$

∎

4 An Approach to Multi-attribute Fuzzy Decision Making

Let $C = \{c_1, c_2, \ldots, c_m\}$ be the attribute set, $A = \{a_1, a_2, \ldots, a_n\}$ be the alternative set. The weight set $W = \{w_1, w_2, \ldots, w_m\}$ is completely unknown with the condition $\sum_{i=1}^{m} w_i = 1$ and $w_i \in [0, 1]$. The evaluation value of a_j to c_i is described by an IVIFS $\tilde{d}_{ij} = \{\langle x_{ij}, [u_A^-(x_{ij}), u_A^+(x_{ij})], [v_A^-(x_{ij}), v_A^+(x_{ij})] \rangle\}$.

Papers [10, 11] pointed out that smaller entropy provides little information, so the smaller entropy of c_i is, the smaller the influence will be exercised on the decision. Then the attribute c_i should be given a smaller weight value. We calculate the weight value by formula [16]:

$$w_i = E_i^{-1} / \sum_{i=1}^{m} E_i^{-1} \tag{5}$$

The concrete steps of the MADM are as follows:

- Step 1 Construct the decision matrix of IVIFS MADM;
- Step 2 Calculate the entropy e_{ij} of every IVIFS based on formula (4), then the entropy of c_i is $E_i = \frac{1}{n}\sum_{j=1}^{n} e_{ij}$;
- Step 3 Calculate the weight value of every attribute based on formula (5);
- Step 4 Let $A^* = \{\langle c_i, [1, 1], [0, 0] \rangle | c_i \in C\}$ be the positive ideal point, calculate the weighted Hamming distance d_j between a_j and the positive ideal point based on formula (2);
- Step 5 Arrange in the order of the value of $d_j, j = 1, 2, \ldots, n$. The smaller value of d_j is, the nearer distance of a_j is to the positive ideal point and the better a_j is.

Numerical example Consider a manufacturer selection problem. The supplier S will choose one manufacturer from a_j $(j = 1, 2, 3)$, which is determined by the evaluation value of five attributes: the quality (c_1), cost (c_2), delivery time (c_3), transportation cost (c_4) and service attitude (c_5). The decision matrix is as follows [9] (Table 1).

Table 1 Decision matrix

c_i/a_j	a_1	a_2	a_3
c_1	$[0.4, 0.5], [0.2, 0.3]$	$[0.3, 0.4], [0.4, 0.5]$	$[0.6, 0.7], [0.1, 0.2]$
c_2	$[0.2, 0.3], [0.5, 0.6]$	$[0.4, 0.5], [0.3, 0.4]$	$[0.3, 0.4], [0.3, 0.5]$
c_3	$[0.4, 0.5], [0.1, 0.3]$	$[0.2, 0.4], [0.3, 0.5]$	$[0.4, 0.6], [0.1, 0.3]$
c_4	$[0.2, 0.4], [0.4, 0.5]$	$[0.4, 0.5], [0.2, 0.3]$	$[0.3, 0.5], [0.3, 0.4]$
c_5	$[0.4, 0.6], [0.1, 0.2]$	$[0.1, 0.2], [0.6, 0.7]$	$[0.4, 0.5], [0.2, 0.3]$

The aggregation result of interval is different with different BUM function, so we choose $\rho(y)$: $y^2, y^3, y, y^{1/2}$, sin $\pi y/2$ to study whether the BUM function affects the decision result.

Step 1 Calculate the entropy value of $c_i(i = 1, 2, \ldots, 5)$ based on formula (4) (Table 2).
Step 2 Calculate the weight value of $c_i(i = 1, 2, \ldots, 5)$ based on formula (5) (Table 3).
Step 3 Calculate the distance d_i between a_j and the positive ideal point based on formula (2) (Table 4).
Step 4 The order of the distance is: $d_3 < d_1 < d_2$. So a_3 is the best, the supplier should choose the third manufacturer, which is same as the result in [9].

Note 4.1 The result shows that different BUM functions (corresponding to different COWA operators) have no influence on the decision making result, which demonstrate the superiority of the above MADM method on another level.

Table 2 The entropy value

Function/E_i	E_1	E_2	E_3	E_4	E_5
y^2	0.4867	0.5163	0.5430	0.5370	0.4874
y^3	0.4913	0.5217	0.5542	0.5438	0.4950
y	0.4783	0.5067	0.5233	0.5250	0.4733
$y^{1/2}$	0.4711	0.4985	0.5075	0.5148	0.4607
sin $\pi y/2$	0.4711	0.4986	0.5072	0.5147	0.4611

Table 3 The weight value

Function/w_i	w_1	w_2	w_3	w_4	w_5
y^2	0.2108	0.1987	0.1890	0.1910	0.2105
y^3	0.2117	0.1993	0.1877	0.1912	0.2101
y	0.2092	0.1975	0.1912	0.1906	0.2144
$y^{1/2}$	0.2079	0.1964	0.1930	0.1902	0.2125
sin $\pi y/2$	0.2079	0.1964	0.1931	0.1902	0.2124

Table 4 The distance between alternative and the positive ideal point

Function/d_j	d_1	d_2	d_3
y^2	0.1262	0.1365	0.1102
y^3	0.1286	0.1384	0.1124
y	0.1215	0.1326	0.1056
$y^{1/2}$	0.1168	0.1286	0.1010
$\sin \pi y/2$	0.1780	0.1812	0.1771

Note 4.2 According to Definitions 3–5, let $f^c(\tilde{A}) = A$, then the Hamming distance and entropy based on COWA operator for IVIFS $\tilde{A}$ can be viewed as the Hamming distance and entropy for IFS A, so we have the following equivalent method to the above one. The concrete steps are as follows:

- Step 1 Construct the decision matrix of IVIFS MADM;
- Step 2 Aggregate the IVIFSs in decision matrix into IFSs, construct the IFS decision matrix;
- Step 3 Calculate the entropy e_{ij} of every IFS by using formula (3), then the entropy of c_i is $E_i = \frac{1}{n}\sum_{j=1}^{n} e_{ij}$;
- Step 3 Calculate the weight value of every attribute by using formula (5);
- Step 4 Let $A^* = \{\langle c_i, 1, 0\rangle | c_i \in C\}$ be the positive ideal point, calculate the weighted Hamming distance d_j between a_j and the positive ideal point by using formula (1);
- Step 5 Arrange in order of the value of $d_j, j = 1, 2, \ldots, n$. The smaller of the value of d_j is, the nearer distance of a_j is to the positive ideal point and the better a_j is.

Practically, the equivalent method is to convert the MADM from IVIFS environment to IFS environment first and then solve it by using the IFS MADM method, which provide us a new strategy. The result, which includes the numerical value of entropy, weight and distance, of the same example solved by the second method is completely consistent with the values in above tables, so we no longer list the computation process.

5 Conclusion

The generalized Hamming distance and entropy for IVIFSs based on the COWA are proposed. We applies the two new measures in IVIFS MADM problems and verifies its effectiveness by an illustrative example. And then we propose the equivalent method of IVIFS MADM which have the same computation result.

References

1. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
2. Atanassov, K., Gargov, G.: Interval-valued intuitionistic fuzzy sets. Fuzzy Sets Syst. **31**(3), 343–349 (1989)
3. Xu, Z.S.: Methods for aggregating interval-valued intuitionistic fuzzy information and their application to decision making. Control Decis. **22**(2), 215–219 (2007)
4. Xu, Z.S., Chen, J.: An approach to group decision making based on interval-valued intuitionistic judgment matrices. Syst. Eng. Theory Pract. **27**(4), 126–133 (2007)
5. Chen, J., Xu, Z.S.: On geometric aggregation over Interval-valued intuitionistic fuzzy information. In: International Conference on Fuzzy Systems & Knowledge Discovery, pp. 466–471 (2007)
6. Ye, J.: Multicriteria fuzzy decision-making method based on a novel accuracy function under interval-valued intuitionistic fuzzy environment. Expert Syst. Appl. **36**(3), 6899–6902 (2009)
7. Li, D.F.: TOPSIS-Based nonlinear-programming methodology for multi-attribute decision making with interval-valued intuitionistic fuzzy sets. IEEE Trans. Syst. **18**(2), 299–311 (2010)
8. Li, D.F.: Linear programming method for MADM with interval-valued intuitionistic fuzzy sets. Expert Syst. Appl. **37**(8), 5939–5945 (2010)
9. Li D.F.: Interval-valued intuitionistic fuzzy set multiattribute decision making methods. In: Li, D.F. (eds.): Intuitionistic Fuzzy Set Decision and Game Analysis Methodologies. National Defence Industry Press, Beijing (2012), pp. 105–132
10. Ye, J.: Multi-criteria fuzzy decision-making method using entropy weights based correlation coefficients of interval-valued intuitionistic fuzzy sets. Appl. Math. Model. **34**(2), 3864–3870 (2010)
11. Gao, Z.H., Wei, C.P.: Formula of interval-valued intuitionistic fuzzy entropy and its applications. Comput. Eng. Appl. **48**(2), 53–55 (2012)
12. Yager, R.: OWA aggregation over a continuous interval argument with applications to decision making. IEEE Trans. Syst. Man Cybern. Part B Cybern. **34**(5), 952–1963 (2004)
13. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst. **114**(3), 505–518 (2000)
14. Szmidt, E., Kacpryzk, J., Bujnowski, P.: How to measure the amount of knowledge conveyed by Atanassov's intuitionistic fuzzy sets. Inf. Sci. **257**, 276–285 (2014)
15. Gao, M.M., Sun, T., Zhu, J.J.: Revised axiomatic definition and structured formula of intuitionistic fuzzy entropy. Control Decis. **29**(3), 470–474 (2014)
16. Wang, C.C., Yao, D.B., Mao, J.J., et al.: Intuitionistic fuzzy multiple attributes decision making method based on entropy and correlation coefficient. J. Comput. Appl. **32**(11), 3002–3004 (2012)

A Novel Weighted Average Index Method of Interval Valued Intuitionistic Fuzzy Sets and Its Application to Outsourced Software Project Risk Assessment

Zhen-hua Zhang, Zhi-jun Li, Xiao-xin Chen, Guo-hua Qu, Yong Hu, Jin-hui Xu and Chao Ma

Abstract A novel weighted average index model of interval valued intuitionistic fuzzy sets (IVIFS) is presented in this paper. By analyzing the interval of membership degree, the interval of non-membership degree and the interval of hesitancy degree, we provide two weighted arithmetic average indexes and two geometric average indexes of IVIFS. And we prove some mathematical properties of these average indexes. Finally, a multiple attribute decision making example applied to outsourced software project risk assessment is given to demonstrate the application of these statistical indexes. The simulation results show that the evaluation methods on these average indexes are all effective.

Keywords Interval valued intuitionistic fuzzy sets · Outsourced software project · Risk assessment · Average index

---

Z. Zhang · Z. Li · X. Chen (✉)
School of Economics and Trade, Guangdong University of Foreign Studies, Guangzhou 510006, China
e-mail: cxxinyannis@hotmail.com

Z. Zhang
e-mail: zhangzhenhua@gdufs.edu.cn

Z. Zhang
Faculty of Business, Environment and Society, Coventry University, Coventry CV1 5FB, UK

G. Qu
School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

Y. Hu · J. Xu
Institute of Big Data and Decision Making, Jinan University, Guangzhou 510632, China

J. Xu · C. Ma
School of Finance, Guangdong University of Foreign Studies, Guangzhou 510006, China

J. Xu
Department of Statistics, Indiana University, Bloomington IN47408, USA

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_26

1 Introduction

In 1965, Zadeh launched fuzzy sets (FS), which has influenced many researchers and has been applied to many application fields, such as pattern recognition, fuzzy reasoning, decision making, etc. In 1980s, Atanassov [1, 2] introduced membership function, non-membership function and hesitancy function, and presented intuitionistic fuzzy sets (IFS) and interval valued intuitionistic fuzzy sets (IVIFS), which generalized the FS theory. In the research field of IFS and IVIFS, Yager [3] discussed its characteristics, more scholars applied it to decision making [4–9] and pattern recognition [10]. Though many scholars studied IFS and IVIFS, few references related to the study of outsourced software project risk assessment based on IFS and IVIFS were proposed. In this paper, we present some novel average index models of IVIFS, and apply them to outsourced software project risk assessment.

First, we introduce the definition of IVIFS and some average indicators of IVIFS. And then, we present two novel average indexes of IVIFS. Finally, we apply the conventional indicators and the novel average indexes to outsourced software project risk assessment. The simulation results show that the method introduced in this paper is an effective method.

2 Conventional Weighted Indicators of IVIFS

Definition 1 An IVIFS A in universe X is given by the following formula [2]:

$$A = \{\langle x, M_A(x), N_A(x)\rangle | x \in X\}. \tag{1}$$

where $M_A(x)$: $X \rightarrow [0, 1]$, $N_A(x)$: $X \rightarrow [0, 1]$ with the condition:

$$\forall x \in X, M_A(x) = [u_A^-(x), u_A^+(x)] \subseteq [0, 1], N_A(x) = [v_A^-(x), v_A^+(x)] \subseteq [0, 1], u_A^+(x) + v_A^+(x) \leq 1.$$

The numbers $M_A(x) \in [0, 1]$, $N_A(x) \in [0, 1]$ denote the interval of membership degree and the interval of non-membership degree of x to A, respectively.

Suppose that $\pi_A^-(x) = 1 - u_A^+(x) - v_A^+(x) \in [0, 1], \pi_A^+(x) = 1 - u_A^-(x) - v_A^-(x) \in [0, 1]$, and we define the interval of hesitant degree $H_A(x) = [\pi_A^-(x), \pi_A^+(x)]$, and then we get IVIFS.

Definition 2 A and B are two IVIFSs over X. For each x ∈ X, we obtain:

$$A \subseteq B \text{ iff } M_A(x) \leq M_B(x), N_A(x) \geq N_B(x),$$
$$M_A(x) \leq M_B(x) \Leftrightarrow \mu_A^-(x) \leq \mu_B^-(x), \mu_A^+(x) \leq \mu_B^+(x),$$
$$N_A(x) \geq N_B(x) \Leftrightarrow \nu_A^-(x) \geq \nu_B^-(x), \nu_A^+(x) \geq \nu_B^+(x).$$

According to IVIFS [2], we define a weighted arithmetic mean on the interval of membership degree and a weighted arithmetic mean on the interval of non-membership degree, respectively:

$$I_M(A) = \sum_{x \in X} w_A(x)(\mu_A^+(x) + \mu_A^-(x)). \tag{2}$$

$$I_{NM}(A) = \sum_{x \in X} w_A(x)(\nu_A^+(x) + \nu_A^-(x)). \tag{3}$$

Based on a dominant ranking function [4], a weighted arithmetic mean on dominant ranking function can be expressed as follows:

$$I_{CT}(A) = \sum_{x \in X} w_A(x)((\mu_A^+(x) - \nu_A^+(x)) + (\mu_A^-(x) - \nu_A^-(x))). \tag{4}$$

Derived from Hong and Choi [5], the following weighted arithmetic mean can be achieved:

$$I_{HC}(A) = \sum_{x \in X} w_A(x)((\mu_A^+(x) + \nu_A^+(x)) + (\mu_A^-(x) + \nu_A^-(x))). \tag{5}$$

where $\mu_A^+(x)$ and $\mu_A^-(x)$ are the degree of membership function, and $\nu_A^+(x)$ and $\nu_A^-(x)$ the degree of non-membership function.

Xu presented the following formula (6) [6, 11]:

$$R_{Xu}(A) = \frac{m(A^+, A)}{m(A^+, A) + m(A^-, A)}. \tag{6}$$

where

$$A^+ = \{ < x, [\max_{A \in \Omega_A}(\mu_A^-(x)), \max_{A \in \Omega_A}(\mu_A^+(x))], [\min_{A \in \Omega_A}(\nu_A^-(x)), \min_{A \in \Omega_A}(\nu_A^+(x))], | x \in X\},$$
$$A^- = \{ < x, [\min_{A \in \Omega_A}(\mu_A^-(x)), \min_{A \in \Omega_A}(\mu_A^+(x))], [\max_{A \in \Omega_A}(\nu_A^-(x)), \max_{A \in \Omega_A}(\nu_A^+(x))], | x \in X\}.$$

Using four distance measures, Xu provided four models from formula (6).

3 Some Average Indexes of IVIFS and Their Properties

According to Xu's formula (6), we define a basic index for each variable $x \in X$.

Definition 3 Suppose that T and F are two types of extreme IVIFSs in X, where $T = \{\langle x, [1, 1], [0, 0] \rangle \mid x \in X\}$ means $M_T(x) = [1, 1]$ and $N_T(x) = [0, 0]$ and

$F = \{\langle x, [0, 0], [1, 1] \rangle \mid x \in X\}$ means $M_F(x) = [0, 0]$ and $N_F(x) = [1, 1]$. We note $I_k(A(x))$ (k = 2, 3) to be an index of IVIFS A for each $x \in X$. And we define:

$$\begin{aligned} I_3(A(x)) &= \frac{m(A(x),F(x))}{m(A(x),F(x))+m(A(x),T(x))} \\ &= \sqrt[p]{|\mu_A^-(x)|^p+|\nu_A^-(x)-1|^p+|\pi_A^-(x)|^p+|\mu_A^+(x)|^p+|\nu_A^+(x)-1|^p+|\pi_A^+(x)|^p} \div \\ &(\sqrt[p]{|\mu_A^-(x)|^p+|\nu_A^-(x)-1|^p+|\pi_A^-(x)|^p+|\mu_A^+(x)|^p+|\nu_A^+(x)-1|^p+|\pi_A^+(x)|^p} \\ &\quad + \sqrt[p]{|\mu_A^-(x)-1|^p+|\nu_A^-(x)|^p+|\pi_A^-(x)|^p+|\mu_A^+(x)-1|^p+|\nu_A^+(x)|^p+|\pi_A^+(x)|^p}). \end{aligned} \tag{7}$$

$$\begin{aligned} I_2(A(x)) &= \frac{m(A(x),F(x))}{m(A(x),F(x))+m(A(x),T(x))} \\ &= \sqrt[p]{|\mu_A^-(x)|^p+|\nu_A^-(x)-1|^p+|\mu_A^+(x)|^p+|\nu_A^+(x)-1|^p} \div \\ &(\sqrt[p]{|\mu_A^-(x)|^p+|\nu_A^-(x)-1|^p+|\mu_A^+(x)|^p+|\nu_A^+(x)-1|^p} \\ &\quad + \sqrt[p]{|\mu_A^-(x)-1|^p+|\nu_A^-(x)|^p+|\mu_A^+(x)-1|^p+|\nu_A^+(x)|^p}). \end{aligned} \tag{8}$$

Then we define the following weighted arithmetic mean index as formula (9) and the following geometric mean index as formula (10):

$$IAM_k(A) = \sum_{x\in X} w_A(x) I_k(A(x)), k = 2, 3. \tag{9}$$

$$IGM_k(A) = \prod_{x\in X} (I_k(A(x)))^{w_A(x)}, k = 2, 3. \tag{10}$$

where m $(A(x), T(x))$ and m $(A(x), F(x))$ are distance measures. And we have $\sum_{x\in X} w_A(x) = 1, w_A(x) \geq 0$. When p = 1, formulas (9) and (10) are based on Hamming distance. Let p = 1, we have (11)–(14):

$$IAM_3(A) = \sum_{x\in X} w_A(x)\left(\frac{2-\nu_A^-(x)-\nu_A^+(x)}{4-\mu_A^-(x)-\mu_A^+(x)-\nu_A^-(x)-\nu_A^+(x)}\right). \tag{11}$$

$$IAM_2(A) = \sum_{x\in X} w_A(x)\left(\frac{2+\mu_A^-(x)+\mu_A^+(x)-\nu_A^-(x)-\nu_A^+(x)}{4}\right). \tag{12}$$

$$IGM_3(A) = \prod_{x\in X} \left(\frac{2-\nu_A^-(x)-\nu_A^+(x)}{4-\mu_A^-(x)-\mu_A^+(x)-\nu_A^-(x)-\nu_A^+(x)}\right)^{w_A(x)}. \tag{13}$$

$$IGM_2(A) = \prod_{x \in X} \left(\frac{2 + \mu_A^-(x) + \mu_A^+(x) - \nu_A^-(x) - \nu_A^+(x)}{4}\right)^{w_A(x)}. \quad (14)$$

We have $0 \leq IAM_k(A) \leq 1$ and $0 \leq IGM_k(A) \leq 1$ for each k and for each A. F indicates that all the example data are the firm opposition party of event A, thus we have $M_F(x) = [0, 0]$, $N_F(x) = [1]$, and $H_A(x) = [0, 0]$. And then we get $IAM_k(F) = 0$ and $IGM_k(F) = 0$, which means that the index of F is zero and the result of F is the worst. Similarly, we have $IAM_k(T) = 1$ and $IGM_k(T) = 1$, which means that the result of T is perfect.

Definition 4 A and B are two IVIFSs over X. For each $x \in X$, we obtain:

$$\begin{aligned}
&A \subseteq B \text{ iff } M_A(x) \leq M_B(x), N_A(x) \geq N_B(x),\\
&M_A(x) \leq M_B(x) \Leftrightarrow \mu_A^-(x) \leq \mu_B^-(x), \mu_A^+(x) \leq \mu_B^+(x),\\
&N_A(x) \geq N_B(x) \Leftrightarrow \nu_A^-(x) \geq \nu_B^-(x), \nu_A^+(x) \geq \nu_B^+(x).
\end{aligned}$$

Theorem 1 *A and B are two IVIFSs, and then we have: If A⊆B then we have* $IAM_2(A) \leq IAM_2(B)$.

Proof

$$\begin{aligned}
&A \subseteq B \leftrightarrow M_A(x) \leq M_B(x), N_A(x) \geq N_B(x),\\
&\rightarrow \begin{cases} \mu_A^-(x) \leq \mu_B^-(x) \\ \mu_A^+(x) \leq \mu_B^+(x) \\ \nu_A^-(x) \geq \nu_B^-(x) \\ \nu_A^+(x) \geq \nu_B^+(x) \end{cases} \rightarrow \begin{cases} 1-\mu_A^-(x) \geq 1-\mu_B^-(x) \\ 1-\mu_A^+(x) \geq 1-\mu_B^+(x) \\ 1-\nu_A^-(x) \leq 1-\nu_B^-(x) \\ 1-\nu_A^+(x) \leq 1-\nu_B^+(x) \end{cases}\\
&\rightarrow \begin{cases} |\mu_A^-(x)|^p \leq |\mu_B^-(x)|^p, |1-\mu_A^-(x)|^p \geq |1-\mu_B^-(x)|^p \\ |\mu_A^+(x)|^p \leq |\mu_B^+(x)|^p, |1-\mu_A^+(x)|^p \geq |1-\mu_B^+(x)|^p \\ |\nu_A^-(x)|^p \geq |\nu_B^-(x)|^p, |1-\nu_A^-(x)|^p \leq |1-\nu_B^-(x)|^p \\ \nu_A^+(x) \geq \nu_B^+(x), |1-\nu_A^+(x)|^p \leq |1-\nu_B^+(x)|^p \end{cases}
\end{aligned}$$

So we have: $IAM_2(A) \leq IAM_2(B)$. And we also have:

Theorem 2 *A and B are two IVIFSs. Assume that A⊆B, and that*

$$\begin{aligned}
&S_{MA} = |\mu_A^-(x)|^p + |1-\nu_A^-(x)|^p + |\mu_A^+(x)|^p + |1-\nu_A^+(x)|^p,\\
&S_{NA} = |1-\mu_A^-(x)|^p + |\nu_A^-(x)|^p + |1-\mu_A^+(x)|^p + |\nu_A^+(x)|^p, S_{HA} = |\pi_A^-(x)|^p + |\pi_A^+(x)|^p,\\
&S_{MB} = |\mu_B^-(x)|^p + |1-\nu_B^-(x)|^p + |\mu_B^+(x)|^p + |1-\nu_B^+(x)|^p,\\
&S_{NB} = |1-\mu_B^-(x)|^p + |\nu_B^-(x)|^p + |1-\mu_B^+(x)|^p + |\nu_B^+(x)|^p, S_{HB} = |\pi_B^-(x)|^p + |\pi_B^+(x)|^p.
\end{aligned}$$

And then we have:

$$S_{MB} \leq S_{NB} \,\&\, S_{HA} \leq S_{HB} \rightarrow IAM_3(A) \leq IAM_3(B),$$
$$S_{MB} \geq S_{NB} \,\&\, S_{HA} \geq S_{HB} \rightarrow IAM_3(A) \leq IAM_3(B).$$

Proof According to definitions of $S_{MA}, S_{NA}, S_{HA}, S_{MB}, S_{NB}, S_{HB}$, we have:

$$IAM_k(A) = \sum_{x \in X} w_A(x) I_3(A(x)) = \frac{\sqrt[p]{S_{MA} + S_{HA}}}{\sqrt[p]{S_{MA} + S_{HA}} + \sqrt[p]{S_{NA} + S_{HA}}} = \frac{1}{1 + \sqrt[p]{\frac{S_{NA} + S_{HA}}{S_{MA} + S_{HA}}}},$$
$$IAM_k(B) = \sum_{x \in X} w_A(x) I_3(B(x)) = \frac{\sqrt[p]{S_{MB} + S_{HB}}}{\sqrt[p]{S_{MB} + S_{HB}} + \sqrt[p]{S_{NB} + S_{HB}}} = \frac{1}{1 + \sqrt[p]{\frac{S_{NB} + S_{HB}}{S_{MB} + S_{HB}}}}.$$

Thus, $IAM_k(A) \leq IAM_k(B) \leftrightarrow \frac{S_{NA} + S_{HA}}{S_{MA} + S_{HA}} \geq \frac{S_{NB} + S_{HB}}{S_{MB} + S_{HB}}$.
From $A \subseteq B$, we have $A \subseteq B \leftrightarrow M_A(x) \leq M_B(x), N_A(x) \geq N_B(x)$,

$$\rightarrow \begin{cases} \mu_A^-(x) \leq \mu_B^-(x) \\ \mu_A^+(x) \leq \mu_B^+(x) \\ \nu_A^-(x) \geq \nu_B^-(x) \\ \nu_A^+(x) \geq \nu_B^+(x) \end{cases} \rightarrow \begin{cases} 1 - \mu_A^-(x) \geq 1 - \mu_B^-(x) \\ 1 - \mu_A^+(x) \geq 1 - \mu_B^+(x) \\ 1 - \nu_A^-(x) \leq 1 - \nu_B^-(x) \\ 1 - \nu_A^+(x) \leq 1 - \nu_B^+(x) \end{cases}$$
$$\rightarrow \begin{cases} |\mu_A^-(x)|^p \leq |\mu_B^-(x)|^p, |1 - \mu_A^-(x)|^p \geq |1 - \mu_B^-(x)|^p \\ |\mu_A^+(x)|^p \leq |\mu_B^+(x)|^p, |1 - \mu_A^+(x)|^p \geq |1 - \mu_B^+(x)|^p \\ |\nu_A^-(x)|^p \geq |\nu_B^-(x)|^p, |1 - \nu_A^-(x)|^p \leq |1 - \nu_B^-(x)|^p \\ \nu_A^+(x) \geq \nu_B^+(x), |1 - \nu_A^+(x)|^p \leq |1 - \nu_B^+(x)|^p \end{cases} \rightarrow \begin{cases} S_{MA} \leq S_{MB} \\ S_{NA} \geq S_{NB} \end{cases}$$

Thus we obtain:

$$S_{MB} \geq S_{NB} \,\&\, S_{HA} \geq S_{HB} \,\&\, \begin{cases} S_{MA} \leq S_{MB} \\ S_{NA} \geq S_{NB} \end{cases}$$
$$\rightarrow \begin{cases} \frac{S_{NB} + S_{HB}}{S_{MB} + S_{HB}} \leq \frac{S_{NB} + S_{HA}}{S_{MB} + S_{HA}} \leq 1 \\ \frac{S_{NB} + S_{HB}}{S_{MB} + S_{HB}} \leq \frac{S_{NB} + S_{HA}}{S_{MB} + S_{HA}} \leq \frac{S_{NA} + S_{HA}}{S_{MA} + S_{HA}} \end{cases} \rightarrow IAM_3(A) \leq IAM_3(B).$$
$$S_{MB} \leq S_{NB} \,\&\, S_{HA} \leq S_{HB} \,\&\, \begin{cases} S_{MA} \leq S_{MB} \\ S_{NA} \geq S_{NB} \end{cases}$$
$$\rightarrow 1 \leq \frac{S_{NB} + S_{HB}}{S_{MB} + S_{HB}} \leq \frac{S_{NB} + S_{HA}}{S_{MB} + S_{HA}} \leq \frac{S_{NA} + S_{HA}}{S_{MA} + S_{HA}} \rightarrow IAM_3(A) \leq IAM_3(B).$$

We draw a conclusion: For each A and each B, $IAM_2(A) \leq IAM_2(B)$. If $A \subseteq B$ and $S_{MB} \leq S_{NB} \,\&\, S_{HA} \leq S_{HB}$ or $A \subseteq B$ and $S_{MB} \geq S_{NB} \,\&\, S_{HA} \geq S_{HB}$, then we obtain $IAM_3(A) \leq IAM_3(B)$.

Theorem 3 A and B are two IVIFSs, and then we have: If $A \subseteq B$ *then we have:*

$$IGM_2(A) \leq IGM_2(B).$$

Proof

$$\begin{aligned}
&A \subseteq B \leftrightarrow M_A(x) \leq M_B(x), N_A(x) \geq N_B(x), \\
&\rightarrow \mu_A^+(x) - \nu_A^+(x) \leq \mu_B^+(x) - \nu_B^+(x), \mu_A^-(x) - \nu_A^-(x) \leq \mu_B^-(x) - \nu_B^-(x). \\
&A \subseteq B \leftrightarrow 1 - \mu_A^-(x) \geq 1 - \mu_B^-(x), 1 - \nu_A^-(x) \leq 1 - \nu_B^-(x); \\
&1 - \mu_A^+(x) \geq 1 - \mu_B^+(x), 1 - \nu_A^+(x) \leq 1 - \nu_B^+(x).
\end{aligned}$$

Therefore we have: $IGM_2(A) \leq IGM_2(B)$.

4 Methodology

Next we will introduce the methodology on the application of the average indexes based on IVIFS above to outsourced software project risk assessment. Considering the specialty of the outsourced software project, we use three first-level attributes to make decision according to Refs. [12–26]: project complexity risks, contractor risks, and customer support and collaboration risks.

Example 1 A manager wants to assess the outsourced software project risk in the process of software development. Given A_i, (i = 1, 2, 3, 4, 5) should be sorted. Assume that three attributes C_1 (contractor risks), C_2 (customer support and collaboration risks), and C_3 (project complexity risks) are taken into consideration, the weight vector of the attributes $C_j (j = 1, 2, 3)$ is $w = (0.5, 0.3, 0.2)^T$. Suppose that the data and the characteristics of the options $A_i (i = 1, 2, 3, 4, 5)$ are shown by IVIFS as follows:

$$\begin{aligned}
A_1 &= \{\langle C_1, [0.6, 0.7], [0, 0.1]\rangle, \langle C_2, [0.1, 0.2], [0.3, 0.4]\rangle, \langle C_3, [0.5, 0.6], \}[0.2, 0.3]\rangle\}, \\
A_2 &= \{\langle C_1, [0.4, 0.5], [0.1, 0.2]\rangle, \langle C_2, [0.3, 0.4], [0.1, 0.2]\rangle, \langle C_3, [0.7, 0.8], \}[0, 0.1]\rangle\}, \\
A_3 &= \{\langle C_1, [0.5, 0.6], [0.1, 0.2]\rangle, \langle C_2, [0.4, 0.5], [0.3, 0.4]\rangle, \langle C_3, [0.8, 0.9], \}[0, 0]\rangle\}, \\
A_4 &= \{\langle C_1, [0.7, 0.8], [0, 0.1]\rangle, \langle C_2, [0.2, 0.3], [0.4, 0.5]\rangle, \langle C_3, [0.6, 0.7], \}[0.1, 0.2]\rangle\}, \\
A_5 &= \{\langle C_1, [0.6, 0.7], [0, 0]\rangle, \langle C_2, [0.7, 0.8], [0.1, 0.2]\rangle, \langle C_3, [0, 0.1], \}[0.5, 0.6]\rangle\}.
\end{aligned}$$

From formulas (11), we obtain the results as follows:

$$IAM_3(A_1) = 0.5 \times \frac{2-0-0.1}{4-0.6-0.7-0-0.1} + 0.3 \times \frac{2-0.3-0.4}{4-0.1-0.2-0.3-0.4} + 0.2 \times \frac{2-0.2-0.3}{4-0.5-0.6-0.2-0.3} \approx 2.543.$$

Similarly, we get the following Table 1.

$I_M(A_3) = I_M(A_4) > I_M(A_5) > I_M(A_1) = I_M(A_2)$, $I_{NM}(A_2) < I_{NM}(A_5) < I_{NM}(A_1) = I_{NM}(A_3) < I_{NM}(A_4)$, thus we get $A_2 \succ A_1, A_5 \succ A_1$, $A_3 \succ A_1$, and $A_3 \succ A_4$. For example, from the membership degree $I_M(A_5) > I_M(A_1) = I_M(A_2)$ and the non-membership degree $I_{NM}(A_2) < I_{NM}(A_5) < I_{NM}(A_1)$, we obtain $A_5 \succ A_1, A_2 \succ A_1$. Similarly, we have $A_3 \succ A_1, A_3 \succ A_4$. Hence, the optimal decision-making is from set $\{A_2, A_3, A_5\}$.

From Table 1, I_{HC} and I_M don't satisfy $A_2 \succ A_1, A_3 \succ A_4$, and IAM_3 doesn't satisfy $A_3 \succ A_4$, I_{NM}, I_{CT}, IAM_2, IGM_2 and IGM_3 satisfy all four conditions on membership degree and non-membership degree.

If the data are given as follows:

$$A_1 = \{\langle C_1, [0.2, 0.2], [0.4, 0.4]\rangle, \langle C_2, [0.7, 0.7], [0.1, 0.1]\rangle, \langle C_3, [0.6, 0.6], \}[0.3, 0.3]\rangle\},$$
$$A_2 = \{\langle C_1, [0.4, 0.4], [0.2, 0.2]\rangle, \langle C_2, [0.5, 0.5], [0.2, 0.2]\rangle, \langle C_3, [0.8, 0.8], \}[0.1, 0.1]\rangle\},$$
$$A_3 = \{\langle C_1, [0.5, 0.5], [0.4, 0.4]\rangle, \langle C_2, [0.6, 0.6], [0.2, 0.2]\rangle, \langle C_3, [0.9, 0.9], \}[0, 0]\rangle\},$$
$$A_4 = \{\langle C_1, [0.3, 0.3], [0.5, 0.5]\rangle, \langle C_2, [0.8, 0.8], [0.1, 0.1]\rangle, \langle C_3, [0.7, 0.7], \}[0.2, 0.2]\rangle\},$$
$$A_5 = \{\langle C_1, [0.8, 0.8], [0.2, 0.2]\rangle, \langle C_2, [0.7, 0.7], [0, 0]\rangle, \langle C_3, [0.1, 0.1], \}[0.6, 0.6]\rangle\}.$$

We have $\mu^-_{A_k}(x) = \mu^+_{A_k}(x), \nu^-_{A_k}(x) = \nu^+_{A_k}(x), \pi^-_{A_k}(x) = \pi^+_{A_k}(x)$, then IVIFS will become IFS, the results will be similar to that from Refs. [7, 8].

Table 1 Evaluation results based on some indicators of IVIFS

Indicators	A_1	A_2	A_3	A_4	A_5	Decision-making
IAM_3	0.620	0.632	0.663	0.664	**0.666**	$A_5 \succ A_4 \succ A_3 \succ A_2 \succ A_1$
IAM_2	0.650	0.675	0.700	0.695	**0.7025**	$A_5 \succ A_3 \succ A_4 \succ A_2 \succ A_1$
IGM_3	2.543	2.577	**2.613**	2.596	2.583	$A_3 \succ A_4 \succ A_5 \succ A_2 \succ A_1$
IGM_2	2.572	2.632	**2.657**	2.626	2.601	$A_3 \succ A_2 \succ A_4 \succ A_5 \succ A_1$
I_M	0.480	0.480	**0.580**	**0.580**	0.560	$A_3 = A_4 \succ A_5 \succ A_2 = A_1$
I_{NM}	0.180	**0.130**	0.180	0.190	0.155	$A_2 = A_5 \succ A_3 \succ A_1 \succ A_4$
I_{CT}	0.300	0.350	0.400	0.390	**0.405**	$A_5 \succ A_3 \succ A_4 \succ A_2 \succ A_1$
I_{HC}	0.660	0.610	0.760	**0.770**	0.715	$A_4 \succ A_3 \succ A_5 \succ A_1 \succ A_2$

5 Application to Outsourced Software Project Risk Assessment

In the following, we will apply the average indexes of IVIFS above to outsourced software project risk assessment.

We design the risks assessment process framework as follows:

(1) Step 1: Attributes selection.
We use three first-level condition attributes: project complexity risks, contractor risks, and customer support and collaboration risks. And all the second-level condition attributes are shown in Table 2.
The decision attribute is Target attribute, including 8 output attributes: Function, Performance, Information Quality, Maintainability, Satisfaction of Customer and User, Company Profits, Completion Degree in Time, Completion Degree in Budget [21, 22]. All the answers are Yes or No. If and only if 8 output attribute values are all Yes then the software project is successful.
In our previous research (Refs. [23–26]), we have presented the attribute framework of outsourced software project risk analysis, in which we use Bayesian networks to set up risk assessment model.

Table 2 Framework of outsourced software project risk analysis

Project complexity risks	References	Customer risks (Support and collaboration risks)	References
1 Estimated cost	[13]	1 Client team collaboration	[14, 15, 18]
2 Lines (KLOC)	[17]	2 Top management support	[14, 15, 18]
3 Number of team members	[18]	3 Client department support	[14, 15, 18]
4 Estimated time	[17]	4 Client development experiment	[14, 18]
5 Technology complexity	[13]	5 Business environment	[14]
6 Fun point	[17]	6 Level of IT application	[15]
7 Real-time and security	[17]	7 Business process	[14]
8 Requirement stability	[12, 15, 19]	Contractor risks	References
9 Number of collaborators	[14, 15]	1 Project manager	[15]
10 Schedule and budget	[13, 15]	2 Development team	[14–16]
11 Industry experience	[14, 15]	3 Plan and control	[15, 16]
		4 Development and test	[20]
		5 Engineering support	[15, 19]

(2) Step 2: Structural equation modeling.
We use structural equation modeling to select the appropriate condition attributes, and the results are shown as Fig. 1. All the attributes that are not significant will be cancelled, where p-value is 0.05.

(3) Step 3: Weights determined in every step.
According to structural equation modeling, we obtain the effect level between different attributes in Fig. 1.
Thus we define the weight as follows:

$$w_{C_{ij}}(x) = \frac{C_{ij}}{\sum_{j=1}^{n_i} C_{ij}}. \quad (15)$$

$$w_{C_i}(x) = \frac{C_i}{\sum_{i=1}^{3} C_i}. \quad (16)$$

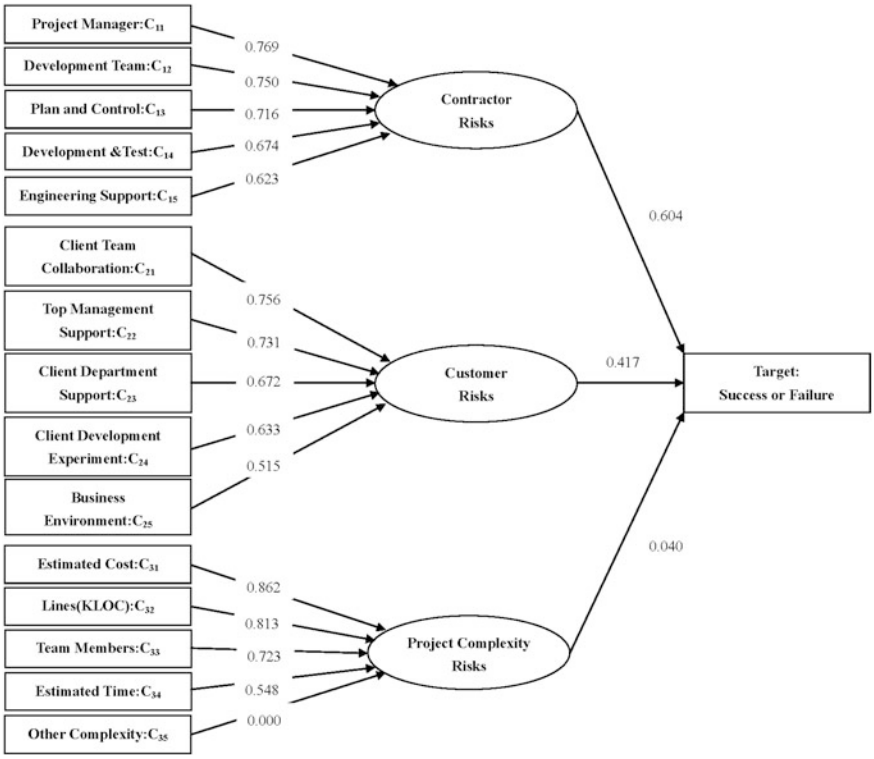


Fig. 1 Effect level of attributes based on structural equation modeling

where C_i means the effect level of first-level condition attribute influencing target. For example, C_1 means the effect level of Contractor Risks (Develop Risks) affecting target, which is 0.604 (in Fig. 1). Similarity, C_{ij} means the effect level of second-level condition attribute in first-level condition attribute. For example, C_{31} means the importance of the Estimated Cost in Project Complexity Risks, which is 0.862.

(4) Step 4: Fuzzy membership and non-membership degree intervals determined. According to five-level survey results, we select triangle module to define membership interval, non-membership Interval, and hesitation interval.

(5) Step 5: Making final decision.
Semi-supervised method: All results are ranked based on the results from formulas (11–13). And we determine a threshold value I_0 according to the success rate of all the outsourced software projects. If the index value $IAM_k(A) > I_0$ then the project is judged as success project, otherwise failure project. Similar to the arithmetic mean model $IAM_k(A)$, the geometric mean model $IGM_k(A)$ can also be designed.
Supervised method: We can also define the threshold value I_0 according to the maximum prediction accuracy. And the assessment standard of supervised method is the same as that of the Semi-supervised method.

6 Experiment Results Analysis

We collect 293 sample data, 260 of them are complete data and 33 incomplete data. In all 260 complete data, 191 of them are success projects and 69 failure projects. And In the experiment, we use 200 to be training sample and the other testing sample. The following Table 3 shows the average accuracy of 10 sampling tests.

It is shown in Table 3 that the prediction effect of the arithmetic average index model is better than that of the geometric average index model. Considering the formulas, we conclude that it is easier for the latter to be influenced than the former by the extreme values. Thus, its decision stability is lower. From Table 3, we also know that the prediction accuracy of supervised model are a little higher than that of semi- supervised model. The predicted results of IGM_3 are the worst, which verifies the conclusion from the theoretical analysis.

Table 3 Average accuracy of IVIFS

Prediction accuracy	Semi-surprised training data	Semi-supervised testing data	Supervised training data	Surprised testing data
IAM_3	0.8865	0.8533	0.899	0.8767
IAM_2	0.9165	0.9083	0.921	0.9183
IGM_3	0.803	0.77	0.8105	0.8017
IGM_2	0.8425	0.8367	0.8795	0.8467

7 Conclusion

We propose a novel average index method of IVIFS derived from Xu's relative average indicators, and apply it to outsourced software project risk assessment. The theoretical analysis show that the arithmetic average index method is better than the geometric average index model, which satisfy all the conditions from the conventional average indicators. And the experiment illustrate the effectiveness of IAM_2 method.

Acknowledgements This paper is funded by the National Natural Science Foundation of China (No. 71271061), the "Twelfth Five-Years" Philosophy and Social Sciences Planning Project of Guangdong Province (No. GD12XGL14), the Science and Technology Innovation Project of Department of Education of Guangdong Province (No. 2013KJCX0072), "Twelfth Five-Years" Philosophy and Social Sciences Planning Project of Guangzhou (No. 14G41), the Natural Science Foundation of Guangdong Province (No. 2014A030313575), the Soft Science Project on Public Research and Capacity Building of Guangdong Province (No. 2015A070704051), the Business Intelligence Key Team of Guangdong University of Foreign Studies (No. TD1202), Student Science and Technology Innovation Cultivating Projects & Climbing Plan Special Key Funds in Guangdong Province (No. 308-GK151011), the Major Education Foundation of Guangdong University of Foreign Studies (No. GYJYZDA14002), the Higher Education Research Project of Guangdong University of Foreign Studies (No. 2016GDJYYJZD004), the National Students Innovation Training Program of China (No. 201511846058).

References

1. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**, 87–96 (1986)
2. Atanassov, K.: Interval valued intuitionistic fuzzy sets. Fuzzy Sets Syst. **31**, 343–349 (1989)
3. Yager, R.R.: Some aspects of intuitionistic fuzzy sets. Fuzzy Optim. Decis. Making **8**, 67–90 (2009)
4. Chen, S.M., Tan, J.M.: Handling multicriteria fuzzy decision-making problems based on vague set theory. Fuzzy Sets Syst. **67**, 163–172 (1994)
5. Hong, D.H., Choi, C.H.: Multicriteria fuzzy decision-making problems based on vague set theory. Fuzzy Sets Syst. **114**, 103–113 (2000)
6. Xu, Z.S.: Some similarity measures of intuitionistic fuzzy sets and their applications to multiple attribute decision making [J]. Fuzzy Optim. Decis. Making **6**, 109–121 (2007)
7. Wei, C.P., Wang, P., Zhang, Y.Z.: Entropy, similarity measure of interval-valued intuitionistic fuzzy sets and their applications. Inf. Sci. **181**, 4273–4286 (2011)
8. Zhang, Z.H., Yang, J.Y., Ye, Y.P., Hu, Y., Zhang, Q.S.: Intuitionistic fuzzy sets with double parameters and its application to dynamic multiple attribute decision making. Inf. Int. Interdisc. J. **15**, 2479–2486 (2012)
9. Zhang, Z.H., Yang, J.Y., Ye, Y.P., Hu, Y., Zhang, Q.S.: A scoring function of intuitionistic fuzzy sets with double parameters and its application to multiple attribute decision making. Inf. Int. Interdisc. J. **15**(11), 4443–4450 (2012)
10. Zhang, Z.H., Wang, M., Hu, Y., Yang, J.Y., Ye, Y.P., Li, Y.F.: A dynamic interval-valued intuitionistic fuzzy sets applied to pattern recognition. Math. Prob. Eng. **2013**(6), 408012, 1–16 (2013)
11. Xia, M.M., Xu, Z.S.: Some new similarity measures for intuitionistic fuzzy values and their application in group decision making. J. Syst. Sci. Syst. Eng. **19**, 430–452 (2010)

12. Jones, C.: Assessment and Control of Software Risks. Prentice Hall, New Jersey (1994)
13. Xu, Z.W., Khoshgoftaar, T.M., Allen, E.B.: Application of fuzzy expert systems in assessing operational risk of software. Inf. Softw. Technol. **45**(7), 373–388 (2003)
14. Xia, W.D., Lee, G.: Complexity of information systems development projects: conceptualization and measurement development. J. Manage. Inf. Syst. **22**(1), 45–84 (2005)
15. Wallace, L., Keil, M.: Software project risks and their effect on outcomes. Commun. ACM **47** (4), 68–73 (2004)
16. Boehm, B.W.: Software risk management: principles and practices. IEEE Softw. **8**(1), 32–41 (1991)
17. Boehm, B.W.: Software Engineering Economics, 1st edn. Prentice-hall, New Jersey (1981)
18. Schmidt, R., Lyytinen, K., Keil, M.: Identifying software project risks: an international Delphi study. J. Manage. Inf. Syst. **17**(4), 5–36 (2001)
19. Karolak, D.W.: Software Engineering Risk Management. IEEE Computer Society Press, Los Alamitos (1996)
20. Jiang, J.J., Klein, G.: An exploration of the relationship between software development process maturity and project performance. Inf. Manage. **41**(3), 279–288 (2004)
21. Wallace, L., Keil, M., Rai, A.: Understanding software project risk: A cluster analysis. Inf. Manage. **42**(1), 115–125 (2004)
22. Nidumolu, S.: The effect of coordination and uncertainty on software project performance: residual performance risk as an intervening variable. Inf. Syst. Res. **6**(3), 191–219 (1995)
23. Hu, Y., Mo, X.Z., Zhang, X.Z., Zeng, Y.R., Du, J.F., Xie, K.: Intelligent analysis model for outsourced software project risk using constraint-based Bayesian network. J. Softw. **7**(2), 440–449 (2012)
24. Hu, Y., Zhang, X.Z., Nagi, E.W.T., Cai, R.C., Liu, M.: Software project risk analysis using Bayesian networks with causality constraints. Decis. Support Syst. **56**(12), 439–449 (2013)
25. Hu, Y., Du, J.F., Zhang, X.Z., Hao, X.L., Nagi, E.W.T., Fan, M., Liu, M.: An integrative framework for intelligent software project risk planning. Decis. Support Syst. **55**(11), 927–937 (2013)
26. Hu, Y., Feng, B., Mo, X.Z., Zhang, X.Z., Nagi, E.W.T., Fan, M., Liu, M.: Cost-Sensitive and ensemble-based prediction model for outsourced software project risk prediction. Decis. Support Syst. **72**(2), 11–23 (2015)

Some Novel Dynamic Fuzzy Sets Models Applied to the Classification of Outsourced Software Project Risk

Zhen-hua Zhang, Guo-hua Qu, Kui-xi Xiao, Yong Hu, Zhi-jun Li, Xiao-xin Chen, Jin-hui Xu and Chao Ma

Abstract Some novel dynamic fuzzy sets (DFS) models, which are the generalization of fuzzy sets (FS) and the dynamization of interval-valued intuitionistic fuzzy sets (IVIFS), are presented in this paper. First, we propose some weighted DFS models from IVIFS. Second, we introduce the distance formula of DFS. Finally, we apply these DFS models and the distance measures to pattern classification of outsourced software project risk to demonstrate the advantages of these DFS models, and the experimental results show that these DFS models are more effective than the conventional clustering algorithms and IVIFS model in pattern classification.

Keywords Interval-valued intuitionistic fuzzy sets · Dynamic fuzzy sets · Pattern classification · Outsourced software project risk

Z. Zhang (✉) · K. Xiao · Z. Li · X. Chen
School of Economics and Trade, Guangdong University of Foreign Studies, Guangzhou 510006, China
e-mail: zhangzhenhua@gdufs.edu.cn; 63221506@qq.com

Z. Zhang
Faculty of Business, Environment and Society, Coventry University, Coventry CV1 5FB, UK

G. Qu
School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

Y. Hu
Institute of Big Data and Decision Making, Jinan University, Guangzhou 510632, China

J. Xu · C. Ma
School of Finance Guangdong University of Foreign Studies, Guangzhou 510006, China

J. Xu
Department of Statistics, Indiana University, Bloomington IN47408, USA

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_27

1 Introduction

In 1965, Professor L.A. Zadeh launched fuzzy sets (FS, [1]). In 1980s, K.T. Atanassov introduced membership function, non-membership function and hesitancy function, and proposed intuitionistic fuzzy sets (IFS, [2]) and interval-valued intuitionistic fuzzy sets (IVIFS, [3]), which generalized the FS theory. Hence, many researchers studied IFS and IVIFS, and applied them to pattern recognition widely [4–13]. However, most of the classic methods are suitable for static model and unsuitable for dynamic model. Taking this into account, Xu [14] presented a dynamic decision making model, which was also studied by Wei [15] and Su et al. [16]. However, traditional decision analysis models on IFS, VS, and IVIFS do not involve the detachment of the absent party. In order to overcome these disadvantages, we present some novel DFS models derived from IVIFS in this paper.

First, we present the definition of DFS and the construction method of DFS. And then, we introduce some DFS models and their distance measures. Finally, we apply these DFS models along with their distance measures to pattern recognition of outsourced software project risk. The simulation results show that these DFS models are more comprehensive and flexible than the IVIFS model. Thus, the model of DFS is valuable for the application of FS to pattern recognition, and it is also useful for the dynamization of fuzzy reasoning, fuzzy decision making, interval-valued intuitionistic fuzzy reasoning and interval-valued intuitionistic fuzzy decision making as in [12, 17, 18].

2 Construction of DFS

Definition 1 An IVIFS A in universe X is given by the following formula [3]:

$$A = \{\langle x,\ M_A(x),\ N_A(x)\rangle | x \in X\}. \tag{1}$$

where $M_A(x) = [u_A^-(x), u_A^+(x)] \subseteq [0,1], N_A(x) = [v_A^-(x), v_A^+(x)] \subseteq [0,1]$ with the condition $u_A^+(x) + v_A^+(x) \leq 1$ for each $x \in X$. The numbers $M_A(x) \in [0,1]$, $N_A(x) \in [0,1]$ denote the interval of membership degree and the interval of non-membership degree of x to A, respectively.

Suppose that $\pi_A^-(x) = 1 - u_A^+(x) - v_A^+(x) \in [0,1], \pi_A^+(x) = 1 - u_A^-(x) - v_A^-(x) \in [0,1]$, and we define the interval of hesitant degree $H_A(x) = [\pi_A^-(x), \pi_A^+(x)]$, and then we get DFS definition as follows.

Definition 2 A DFS A_1^* derived from IVIFS in universe X is denoted by:

$$A_1^* = \{\langle x,\ \mu_{A1}^*(x),\ v_{A1}^*(x)\rangle | x \in X\}. \tag{2}$$

where $\mu^*_{A1}(x) = \mu^-_A(x) + \lambda_1\pi^+_A(x)$ and $v^*_{A1}(x) = v^-_A(x) + (1-\lambda_1)\pi^+_A(x)$ with the condition $\lambda_1 \in [0, 1]$. $\mu^*_{A1}(x)$ and $v^*_{A1}(x)$ are membership function and non-membership function of x to A, respectively.

From Definition 2, let all sample data be divided into three parts, $\mu^-_A(x)$ being the firm support party of event A, $v^-_A(x)$ representing the firm opposition party of event A, and $\pi^+_A(x)$ showing all the absent party that may become either the support party or the opposition party. In the absent party, if there is $\lambda\pi^+_A(x)$ sample supporting event A and $(1-\lambda)\pi^+_A(x)$ sample opposing event A, we have DFS denoted by Definition 2. Obviously, DFS is an extension of FS and a dynamic method of IVIFS.

Similarly, we can define DFS as follows.

Definition 3 A DFS A^*_2 derived from IVIFS in universe X can be also denoted by:

$$A^*_2 = \{\langle x,\ \mu^*_{A2}(x),\ v^*_{A2}(x)\rangle | x \in X\}. \tag{3}$$

where $\mu^*_{A2}(x) = \mu^+_A(x) + \lambda_2\pi^-_A(x)$ and $v^*_{A2}(x) = v^+_A(x) + (1-\lambda_2)\pi^-_A(x)$ with the condition $\lambda_2 \in [0, 1]$. $\mu^*_{A2}(x)$ and $v^*_{A2}(x)$ are membership function and non-membership function of x to A, respectively.

Theorem 1 *Let A be an DFS as mentioned above, then*

$$\mu^*_{A1}(x) + \nu^*_{A1}(x) = \mu^*_{A2}(x) + \nu^*_{A2}(x) = 1.$$

According to Definitions 2 and 3, we have Theorem 1.

From Definitions 2 and 3, we have Definition 4 as follows.

Definition 4 A DFS A^* derived from IVIFS in universe X can be also denoted by:

$$A^* = \{\langle x,\ \mu^*_A(x),\ v^*_A(x)\rangle | x \in X\}. \tag{4}$$

$$\begin{aligned}
\mu^*_A(x) &= \lambda_3(\mu^-_A(x) + \lambda_1\pi^+_A(x)) + \lambda_4(\mu^+_A(x) + \lambda_2\pi^-_A(x)),\\
v^*_A(x) &= \lambda_3[v^-_A(x) + (1-\lambda_1)\pi^+_A(x)] + \lambda_4[v^+_A(x) + (1-\lambda_2)\pi^-_A(x)],\\
&\lambda_i \in [0, 1], i = 1, 2, 3, 4; \lambda_3 + \lambda_4 = 1.
\end{aligned}$$

where λ_3 and λ_4 denote the weight of $\mu^*_{A1}(x)$ and $\mu^*_{A2}(x)$, and $\mu^*_A(x)$ and $v^*_A(x)$ are membership function and non-membership function of x to A, respectively. And we also have $\mu^*_A(x) + \nu^*_A(x) = 1$.

From Definition 4, Definition 5 is more simple than Definition 4.

Definition 5 A DFS A^* derived from IVIFS in universe X can be also denoted by:

$$A^* = \{\langle x,\ \mu^*_A(x),\ v^*_A(x)\rangle | x \in X\}. \tag{5}$$

$$\mu_A^*(x) = \lambda(\mu_A^-(x) + \lambda_1\pi_A^+(x)) + (1-\lambda)(\mu_A^+(x) + \lambda_2\pi_A^-(x)),$$
$$\nu_A^*(x) = \lambda[\nu_A^-(x) + (1-\lambda_1)\pi_A^+(x)] + (1-\lambda)[\nu_A^+(x) + (1-\lambda_2)\pi_A^-(x)],$$
$$\lambda \in [0,1], \lambda_1 \in [0,1], \lambda_2 \in [0,1].$$

where λ and $1-\lambda$ denote the weight of $\mu_{A1}^*(x)$ and $\mu_{A2}^*(x)$, and $\mu_A^*(x)$ and $\nu_A^*(x)$ are membership function and non-membership function of x to A, respectively.

For example, if $\lambda_i \in [0,1], \lambda = 0.5$, then we obtain:

$$\mu_A^*(x) = \frac{\mu_A^-(x) + \mu_A^+(x)}{2} + \frac{\lambda_1\pi_A^+(x) + \lambda_2\pi_A^-(x)}{2},$$
$$\nu_A^*(x) = \frac{\nu_A^-(x) + \nu_A^+(x)}{2} + \frac{(1-\lambda_1)\pi_A^+(x) + (1-\lambda_2)\pi_A^-(x)}{2}.$$

3 Distance Measures

According to the membership function and the non-membership function of DFS, a weighted standardized Minkowski distance can be defined as follows:

$$d_{DFS}(A^*, B^*) = \sqrt[k]{\sum_{x \in X} w_{AB}(x)(C_1|\mu_A^*(x) - \mu_B^*(x)|^k + C_2|\nu_A^*(x) - \nu_B^*(x)|^k)}. \quad (6)$$

where the weight of membership function is C_1 and that of non-membership function is C_2, and $w_{AB}(x)$ denotes the weight of element x, and we have $C_1 \geq 0, C_2 \geq 0, C_1 + C_2 = 1$. $w_{AB}(x) \geq 0, \sum_{x \in X} w_{AB}(x) = 1$.

According to Theorem 1, we have $\mu_A^*(x) + \nu_A^*(x) = 1$, and then we have:

$$\begin{aligned} d_{DFS}(A^*, B^*) &= \sqrt[k]{\sum_{x \in X} w_{AB}(x)(C_1|\mu_A^*(x) - \mu_B^*(x)|^k + (1-C_1)|u_A^*(x) - u_B^*(x)|^k)} \\ &= \sqrt[k]{\sum_{x \in X} w_{AB}(x)|\mu_A^*(x) - \mu_B^*(x)|^k}. \end{aligned} \quad (7)$$

From Definition 2, if $k = 2$, $\lambda \in [0, 1]$, we obtain:

$$d_{DFS}(A^*, B^*) = \sqrt{\begin{aligned} &\sum_{x \in X} w_{AB}(x)\{\lambda^2[(\mu_A^-(x) - \mu_B^-(x))^2 + \lambda_1^2(\pi_A^+(x) - \pi_B^+(x))^2 \\ &+ 2\lambda_1(\mu_A^-(x) - \mu_B^-(x))(\pi_A^+(x) - \pi_B^+(x))] \\ &+ (1-\lambda)^2[(\mu_A^+(x) - \mu_B^+(x))^2 + \lambda_2^2(\pi_A^-(x) - \pi_B^-(x))^2 \\ &+ 2\lambda_2(\mu_A^+(x) - \mu_B^+(x))(\pi_A^-(x) - \pi_B^-(x))]\} \end{aligned}}. \quad (8)$$

From (8), we define a simple weighted standardized distance measure of DFS as follows:

$$d(A^*, B^*) = \sum_{x \in X} w_{AB}(x)\{\lambda^2[(\mu_A^-(x) - \mu_B^-(x))^2 + \lambda_1^2(\pi_A^+(x) - \pi_B^+(x))^2 \\ + 2\lambda_1(\mu_A^-(x) - \mu_B^-(x))(\pi_A^+(x) - \pi_B^+(x))] + (1-\lambda)^2[(\mu_A^+(x) - \mu_B^+(x))^2 \\ + \lambda_2^2(\pi_A^-(x) - \pi_B^-(x))^2 + 2\lambda_2(\mu_A^+(x) - \mu_B^+(x))(\pi_A^-(x) - \pi_B^-(x))]\}. \quad (9)$$

where $\lambda \in [0, 1]$, $w_{AB}(x) \geq 0$, $\sum_{x \in X} w_{AB}(x) = 1$.

4 Application to Classification of Outsourced Software Risk

In the following, a pattern recognition problem about the classification of outsourced software project risk [5–7] is used to illustrate DFS derived from IVIFS and its distance measure above.

Example 1 Assume that there are four classes of outsourced software project risk, which are represented by the IVIFS $A_i = \{\langle x, M_{A_i}(x), N_{A_i}(x)\rangle | x \in X\}$ in the feature space $X = \{x_1, x_2, \ldots, x_{12}\}$ whose weight vector is $w = \{0.1, 0.05, 0.08, 0.06, 0.03, 0.07, 0.09, 0.12, 0.15, 0.07, 0.13, 0.05\}^T$, and there is an unknown outsourced software project B. We aim to justify which class the unknown pattern B belongs to. The following Table 1 show all the attributes we use, which have been proved to be most important attributes between customers and contractors when they develop a outsourced software [19–22].

Table 1 Framework of outsourced software project risk analysis

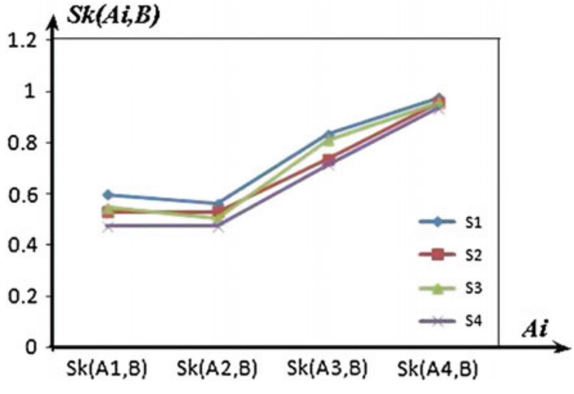
Customer risks (Support and collaboration risks)	References	Contractor Risks	References
1 Client team collaboration (x_6)	[25–27]	1 Project manager (x_1)	[26]
2 Top management support (x_7)	[25–27]	2 Development team (x_2)	[25, 26, 28]
3 Client department support (x_8)	[25–27]	3 Plan and control (x_3)	[26, 28]
4 Client development experiment (x_9)	[25, 27]	4 Development and test (x_4)	[29]
5 Business environment (x_{10})	[25]	5 Engineering support (x_5)	[26, 30]
6 Level of IT application (x_{11})	[26]		
7 Business process (x_{12})	[25]		

$$
\begin{aligned}
A_1 = \{ & \langle x_1,\ [0.1,\ 0.2],\ [0.5,\ 0.6]\rangle, \langle x_2,\ [0.1,\ 0.2],\ [0.7,\ 0.8], \langle x_3,\ [0.5,\ 0.6],\ [0.3,\ 0.4]\rangle, \\
& \langle x_4,\ [0.8,\ 0.9],\ [0,\ 0.1]\rangle, \langle x_5,\ [0.4,\ 0.5],\ [0.3,\ 0.4]\rangle, \langle x_6,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \\
& \langle x_7,\ [0.3,\ 0.4],\ [0.5,\ 0.6]\rangle, \langle x_8,\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_9,\ [0.2,\ 0.3],\ [0.6,\ 0.7]\rangle, \\
& \langle x_{10},\ [0.4,\ 0.5],\ [0.4,\ 0.5]\rangle, \langle x_{11},\ [0.7,\ 0.8],\ [0.1,\ 0.2]\rangle, \langle x_{12},\ [0.4,\ 0.5],\ [0.4,\ 0.5]\rangle \}. \\
A_2 = \{ & \langle x_1,\ [0.5,\ 0.6],\ [0.3,\ 0.4]\rangle, \langle x_2,\ [0.6,\ 0.7],\ [0.1,\ 0.2]\rangle, \langle x_3,\ [1.0,\ 1.0],\ [0,\ 0]\rangle \\
& \langle x_4,\ [0.1, 0.2],\ [0.6,\ 0.7]\rangle, \langle x_5,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \langle x_6,\ [0.7,\ 0.8],\ [0.1,\ 0.2]\rangle, \\
& \langle x_7,\ [0.5,\ 0.6],\ [0.3,\ 0.4]\rangle, \langle x_8,\ [0.6,\ 0.7],\ [0.2,\ 0.3]\rangle, \langle x_9,\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \\
& \langle x_{10},\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_{11},\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \langle x_{12},\ [0.7,\ 0.8],\ [0.1,\ 0.2]\rangle \}. \\
A_3 = \{ & \langle x_1,\ [0.4,\ 0.5],\ [0.3,\ 0.4]\rangle, \langle x_2,\ [0.6,\ 0.7],\ [0.2,\ 0.3]\rangle, \langle x_3,\ [0.9,\ 1.0],\ [0,\ 0]\rangle, \\
& \langle x_4,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \langle x_5,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \langle x_6,\ [0.6,\ 0.7],\ [0.2,\ 0.3]\rangle, \\
& \langle x_7,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_8,\ [0.2,\ 0.3],\ [0.6,\ 0.7]\rangle, \langle x_9,\ [0.5,\ 0.6],\ [0.2,\ 0.4]\rangle, \\
& \langle x_{10},\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_{11},\ [0.3,\ 0.4],\ [0.4,\ 0.5]\rangle, \langle x_{12},\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle \}. \\
A_4 = \{ & \langle x_1,\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_2,\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_3,\ [0.8,\ 0.9],\ [0,\ 0.1]\rangle, \\
& \langle x_4,\ [0.7,\ 0.8],\ [0.1,\ 0.2]\rangle, \langle x_5,\ [0,\ 0.1],\ [0.7,\ 0.9]\rangle, \langle x_6,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \\
& \langle x_7,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_8,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_9,\ [0.4,\ 0.5],\ [0.3,\ 0.4]\rangle, \\
& \langle x_{10},\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_{11},\ [0.3,\ 0.4],\ [0.4,\ 0.5]\rangle, \langle x_{12},\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle \}. \\
B = \{ & \langle x_1,\ [0.9,\ 1.0],\ [0,\ 0]\rangle, \langle x_2,\ [0.9,\ 1.0],\ [0,\ 0]\rangle, \langle x_3,\ [0.7,\ 0.8],\ [0.1,\ 0.2]\rangle, \\
& \langle x_4,\ [0.6,\ 0.7],\ [0.1,\ 0.2]\rangle, \langle x_5,\ [0,\ 0.1],\ [0.8,\ 0.9]\rangle, \langle x_6,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \\
& \langle x_7,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_8,\ [0.1,\ 0.2],\ [0.7,\ 0.8]\rangle, \langle x_9,\ [0.4,\ 0.5],\ [0.3,\ 0.4]\rangle, \\
& \langle x_{10},\ [1.0,\ 1.0],\ [0,\ 0]\rangle, \langle x_{11},\ [0.3,\ 0.4],\ [0.4,\ 0.5]\rangle, \langle x_{12},\ [0,\ 0.1],\ [0.7,\ 0.9]\rangle \}.
\end{aligned}
$$

Xu presented four similarity measures in [6], and obtained the following values on similarity measure and got Fig. 1.

$$
\begin{aligned}
& s_1(A_1, B) = 0.597,\ s_1(A_2, B) = 0.561,\ s_1(A_3, B) = 0.833,\ s_1(A_4, B) = 0.976; \\
& s_2(A_1, B) = 0.530,\ s_2(A_2, B) = 0.529,\ s_2(A_3, B) = 0.734,\ s_2(A_4, B) = 0.951; \\
& s_3(A_1, B) = 0.545,\ s_3(A_2, B) = 0.503,\ s_3(A_3, B) = 0.810,\ s_3(A_4, B) = 0.956; \\
& s_4(A_1, B) = 0.473,\ s_4(A_2, B) = 0.473,\ s_4(A_3, B) = 0.712,\ s_4(A_4, B) = 0.934.
\end{aligned}
$$

Fig. 1 Similarity measure between A_i and B



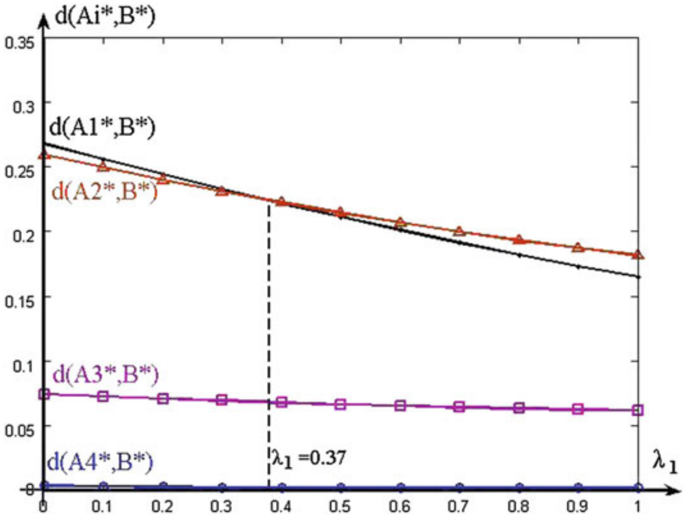


Fig. 2 DFS distance measure for Definition 2

In [6], Xu concluded that the degree of similarity between A_4 and B is the largest one, as is shown in Fig. 2. Therefore, the unknown pattern B should belong to the pattern A_4. In 2011, Wei et al. also demonstrated that B should be A_4 using another similarity measure in [7].

We calculate the DFS distance measure between Ai and B by formula (9), if $\lambda = 1$ then we obtain:

$$d(A_1^*, B^*) = 0.2679 - 0.1244\lambda + 0.0213\lambda^2, d(A_2^*, B^*) = 0.2592 - 0.102\lambda + 0.0243\lambda^2,$$
$$d(A_3^*, B^*) = 0.0745 - 0.019\lambda + 0.0064\lambda^2, d(A_4^*, B^*) = 0.0036 - 0.0042\lambda + 0.0029\lambda^2.$$

From the distance measures above, we have Fig. 2. Obviously, the distance between A_4 and B is the smallest for each $\lambda \in [0, 1]$, thus, B belongs to the pattern A_4.

If $\lambda = 0$ then we obtain the following distance measures and Fig. 3.

$$d(A_1^*, B^*) = 0.2398 - 0.0364\lambda + 0.0077\lambda^2, d(A_2^*, B^*) = 0.2268 - 0.0102\lambda + 0.0033\lambda^2,$$
$$d(A_3^*, B^*) = 0.0745 - 0.0058\lambda + 0.0031\lambda^2, d(A_4^*, B^*) = 0.0021 - 0.0012\lambda + 0.0006\lambda^2.$$

If $\lambda = 0.5$ then we obtain the following distance measures and Fig. 4.

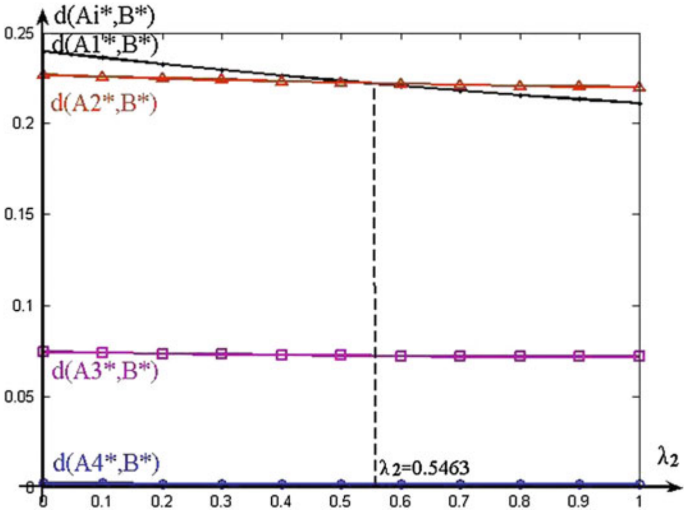


Fig. 3 DFS distance measure for Definition 3

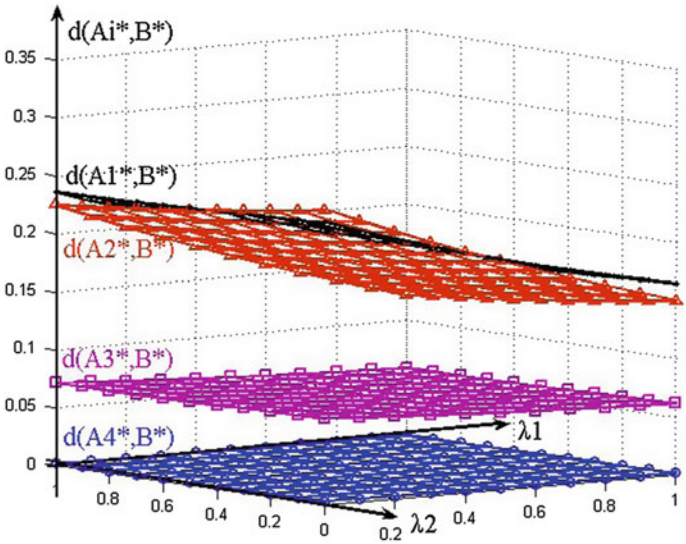


Fig. 4 DFS distance measure for Definition 4

$$4\times d(A_1^*,B^*)=0.25385-0.1244\lambda_1+0.0213\lambda_1^2-0.0364\lambda_2+0.0077\lambda_2^2,$$
$$4\times d(A_2^*,B^*)=0.243-0.102\lambda_1+0.0243\lambda_1^2-0.0102\lambda_2+0.0033\lambda_2^2,$$
$$4\times d(A_3^*,B^*)=0.0745-0.019\lambda_1+0.0064\lambda_1^2-0.0058\lambda_2+0.0031\lambda_2^2,$$
$$4\times d(A_4^*,B^*)=0.00285-0.0042\lambda_1+0.0029\lambda_1^2-0.0012\lambda_2+0.0006\lambda_2^2.$$

For the convenience of the comparison among these models, we take the equations above as the distance measures of DFS from Definition 4.

Compare Fig. 2 with Fig. 3, we conclude that DFS from Definition 3 is more stable than DFS from Definition 2. In theory, we have $0\le\pi_A^-(x)\le\pi_A^+(x)$, which means that the hesitancy degree of DFS from Definition 2 is more than that from Definition 3. Thus, the variation of DFS from Definition 2 is quicker than that from Definition 3, which is also demonstrated in Fig. 4.

According to Figs. 2, 3 and 4, we have

$$d(A_4^*,B)\langle d(A_3^*,B)\langle \min imum(d(A_1^*,B),d(A_2^*,B))$$

for each $\lambda\in[0, 1]$ and for each $\lambda_i\in[0, 1]$, $i = 1, 2$. Thus, the unknown pattern B should belong to the pattern A_4, which is the same as the results in [6, 7].

Though the results in this paper are the same as the results in [6, 7], Xu's similarity measures and Wei's similarity measures are suitable for static model and unsuitable for dynamic model. If the results are different from the practical results, the similarity measures above will not work. However, for DFS method, we can adjust the parameter to appropriate value to deal with this difference. For example, for Xu's method s_k (k = 1, 2, 3), we have $s_k(A_4, B) > s_k(A_3, B) > s_k(A_1, B) > s_k(A_2, B)$, which are the same as the results of DFS from Definition 2 for $\lambda > 0.37$. For Xu's method s_4, we have $s_4(A_4, B) > s_4(A_3, B) > s_4(A_1, B) = s_4(A_2, B)$, which are the same as the results of DFS from Definition 2 for $\lambda = 0.37$. Moreover, if the practical results is $s(A_1, B) > s(A_2, B)$, then Xu's method will not work. However, if $\lambda < 0.37$, the DFS method from Definition 2 is suitable for this case. Similarly, we can analyze the same case from Definitions 3 and 4. Thus, we draw a conclusion that these DFS methods are more comprehensive than some conventional similarity methods of IVIFS.

5 Algorithm Steps for Outsourced Software Risk Classification

We collect 293 sample data, 260 of them are complete data and 33 incomplete data. In all 260 complete data, 191 of them are success projects and 69 failure projects. The decision attribute is Target attribute, including 8 output attributes: Function, Performance, Information Quality, Maintainability, Satisfaction of Customer and

User, Company Profits, Completion Degree in Time, Completion Degree in Budget [23, 24]. All the answers are Yes or No. If and only if 8 output attribute values are all Yes then the software project is successful.

We adopt the following algorithm step:

Step 1 Fuzzification.

We use five scales to measure the difference of attribute value. Therefore, for "Bad" set A, we define:

$$\mu_A^+(x) = 1.2 - 0.2x, \mu_A^-(x) = 1 - 0.2x, \nu_A^+(x) = 0.2x, \mu_A^-(x) = 0.2x - 0.2;$$

For "Common" set A, if $x \leq 3$, we define:

$$M_A(x) = [\mu_A^-(x), \mu_A^+(x)] = [\max\{|0.5x - 0.5|, |0.5x - 0.7|\}, \min\{|0.5x - 0.5|, |0.5x - 0.7|\},$$
$$N_A(x) = [\nu_A^-(x), \nu_A^+(x)] = [\max\{1.5 - 0.5x, 0.5x - 0.2\}, \min\{1.5 - 0.5x, 0.5x - 0.2\}.$$

If $x > 3$, we define:

$$M_A(x) = [\mu_A^-(x), \mu_A^+(x)] = [\max\{|2.5 - 0.5x|, |2.3 - 0.5x|\}, \min\{|2.5 - 0.5x|, |2.3 - 0.5x|\},$$
$$N_A(x) = [\nu_A^-(x), \nu_A^+(x)] = [\max\{0.5x - 1.5, 0.5x - 1.3\}, \min\{0.5x - 1.5, 0.5x - 1.3\}.$$

For "Good" set A, we define:

$$\mu_A^+(x) = 0.2x, \mu_A^-(x) = 0.2x - 0.2, \nu_A^+(x) = 1.2 - 0.2x, \mu_A^-(x) = 1 - 0.2x.$$

Step 2 Weighted fuzzification according to interval value intuitionistic fuzzy sets.

Applying the probability of successful project and failure project to construct fuzzy membership interval and non-membership interval.

Step 3 Based on the result of K-means algorithm to set up 8 category center and then we fuzzify them using interval intuitionistic fuzzy formula above. And we obtain the following category center:

$$A_1 = \{\langle x_1, [0.9, 1], [0, 0.1]\rangle, \langle x_2, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_3, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_4, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_5, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_6, [0.9, 1], [0, 0.1]\rangle, \langle x_7, [0.9, 1], [0, 0.1]\rangle, \langle x_8, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_9, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{10}, [0.9, 1], [0, 0.1]\rangle, \langle x_{11}, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{12}, [0.7, 0.9], [0.1, 0.3]\rangle\}.$$

$$A_2 = \{\langle x_1, [0, 0.2], [0.8, 1]\rangle, \langle x_2, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_3, [0.9, 1], [0, 0.1]\rangle, \langle x_4, [0.9, 1], [0, 0.1]\rangle, \langle x_5, [0.9, 1], [0, 0.1]\rangle, \langle x_6, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_7, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_8, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_9, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{10}, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{11}, [0, 0.2], [0.8, 1]\rangle, \langle x_{12}, [0.9, 1], [0, 0.1]\rangle\}.$$

$$A_3 = \{\langle x_1, [0.3, 0.5], [0.5, 0.7]\rangle, \langle x_2, [0.3, 0.5], [0.5, 0.7]\rangle, \langle x_3, [0.3, 0.5], [0.5, 0.7]\rangle \langle x_4, [0.3, 0.5], [0.5, 0.7]\rangle, \langle x_5, [0.3, 0.5], [0.5, 0.7]\rangle, \langle x_6, [0.2, 0.4], [0.4, 0.6]\rangle, \langle x_7, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_8, [0.2, 0.4], [0.4, 0.6]\rangle, \langle x_9, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{10}, [0.2, 0.4], [0.4, 0.6]\rangle, \langle x_{11}, [0.2, 0.4], [0.4, 0.6]\rangle, \langle x_{12}, [0.2, 0.4], [0.4, 0.6]\rangle\}.$$

$$A_4 = \{\langle x_1, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_2, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_3, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_4, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_5, [0.2, 0.4], [0.4, 0.6]\rangle, \langle x_6, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_7, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_8, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_9, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{10}, [0.9, 1.0], [0, 0.1]\rangle, \langle x_{11}, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{12}, [0.4, 0.6], [0.4, 0.6]\rangle\}.$$

$$A_4 = \{\langle x_1, [1.0, 1.0], [0, 0]\rangle, \langle x_2, [1.0, 1.0], [0, 0]\rangle, \langle x_3, [0.8, 0.9], [0, 0.1]\rangle, \langle x_4, [0.7, 0.8], [0.1, 0.2]\rangle, \langle x_5, [0, 0.1], [0.7, 0.9]\rangle, \langle x_6, [0, 0.1], [0.8, 0.9]\rangle, \langle x_7, [0.1, 0.2], [0.7, 0.8]\rangle, \langle x_8, [0.1, 0.2], [0.7, 0.8]\rangle, \langle x_9, [0.4, 0.5], [0.3, 0.4]\rangle, \langle x_{10}, [1.0, 1.0], [0, 0]\rangle, \langle x_{11}, [0.3, 0.4], [0.4, 0.5]\rangle, \langle x_{12}, [0, 0.1], [0.8, 0.9]\rangle\}.$$

$$A_5 = \{\langle x_1, [0.9, 1], [0, 0.1]\rangle, \langle x_2, [0.9, 1], [0, 0.1]\rangle, \langle x_3, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_4, [0.9, 1], [0, 0.1]\rangle, \langle x_5, [0.9, 1], [0, 0.1]\rangle, \langle x_6, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_7, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_8, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_9, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{10}, [0, 0.1], [0.9, 1]\rangle, \langle x_{11}, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{12}, [0.1, 0.3], [0.7, 0.9]\rangle\}.$$

$$A_6 = \{\langle x_1, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_2, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_3, [0.5, 0.7], [0.3, 0.5]\rangle, \langle x_4, [0.5, 0.7], [0.3, 0.5]\rangle, \langle x_5, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_6, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_7, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_8, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_9, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{10}, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{11}, [0.9, 1], [0, 0.1]\rangle, \langle x_{12}, [0.4, 0.6], [0.4, 0.6]\rangle\}.$$

$$A_7 = \{\langle x_1, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_2, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_3, [0.3, 0.5], [0.5, 0.7]\rangle \langle x_4, [0.1, 0.3], [0.7, 0.9]\rangle, \langle x_5, [0.1, 0.3], [0.7, 0.9]\rangle, \langle x_6, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_7, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_8, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_9, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{10}, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{11}, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{12}, [0.4, 0.6], [0.4, 0.6]\rangle\}.$$

$$A_8 = \{\langle x_1, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_2, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_3, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_4, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_5, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_6, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_7, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_8, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_9, [0.4, 0.6], [0.4, 0.6]\rangle, \langle x_{10}, [0.7, 0.9], [0.1, 0.3]\rangle, \langle x_{11}, [0.1, 0.3], [0.7, 0.9]\rangle, \langle x_{12}, [0.4, 0.6], [0.4, 0.6]\rangle\}$$

Step 4 Assume that step length change from 0.01 to 0.1, calculate the distance between all the training sample and testing sample, and then we determine the classification.

Table 2 Average accuracy of IVIFS

Prediction accuracy / Prediction method	Failure project	Success project	Total project
DFS ($\lambda = 0.06$)	**0.916**	**0.667**	**0.850**
IVIFS	**0.932**	**0.478**	**0.812**
K-means clustering	0.906	0.435	0.781
Fuzzy clustering	0.901	0.406	0.769
CHAID decision tree	0.859	0.638	0.800
FNN	0.780	0.710	0.762
BPNN	0.963	0.188	0.758

6 Experiment Results and Conclusion

From Table 2, we conclude that the pattern classification of IVIFS and DFS are more effective than traditional clustering method.

In this paper, we propose some novel dynamic fuzzy sets models derived from IVIFS, which not only involve membership function and non-membership function, but also involve the detachment of hesitancy function. Therefore, they are more comprehensive and flexible than the IVIFS model.

Acknowledgements This paper is funded by the National Natural Science Foundation of China (No. 71271061), the "Twelfth Five-Years" Philosophy and Social Sciences Planning Project of Guangdong Province (No. GD12XGL14), the Science and Technology Innovation Project of Department of Education of Guangdong Province (No. 2013KJCX0072), "Twelfth Five-Years" Philosophy and Social Sciences Planning Project of Guangzhou (No. 14G41), the Natural Science Foundation of Guangdong Province (No. 2014A030313575), the Soft Science Project on Public Research and Capacity Building of Guangdong Province (No. 2015A070704051), the Business Intelligence Key Team of Guangdong University of Foreign Studies (No. TD1202), Student Science and Technology Innovation Cultivating Projects & Climbing Plan Special Key Funds in Guangdong Province (No. 308-GK151011), the Major Education Foundation of Guangdong University of Foreign Studies (No. GYJYZDA14002), the Higher Education Research Project of Guangdong University of Foreign Studies (No. 2016GDJYYJZD004), the National Students Innovation Training Program of China (No. 201511846058).

References

1. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
2. Atanassov, k: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**, 87–96 (1986)
3. Atanassov, k: Interval valued intuitionistic fuzzy sets. Fuzzy Sets Syst. **31**, 343–349 (1989)
4. Li, D.F., Cheng, C.T.: New similarity measures of intuitionistic fuzzy sets and applications to pattern recognitions. Pattern Recogn. Lett. **23**, 221–225 (2002)
5. Wang, W.Q., Xin, X.L.: Distances measure between intuitionistic fuzzy sets. Pattern Recogn. Lett. **26**, 2063–2069 (2005)

6. Xu, Z.S.: On similarity measures of interval-valued intuitionistic fuzzy sets and their application to pattern recognitions. J. Southeast Univ. (English Edition) **23**, 139–143 (2007)
7. Wei, C.P., Wang, P., Zhang, Y.Z.: Entropy, similarity measure of interval-valued intuitionistic fuzzy sets and their applications. Inf. Sci. **181**, 4273–4286 (2011)
8. Wan, S.P.: Applying interval-value vague set for multi-sensor target recognition. Int. J. Innovative Comput. Inf. Control **7**, 955–963 (2011)
9. Xu, Z.S.: A method based on distance measure for interval-valued intuitionistic fuzzy group decision making. Inf. Sci. **180**, 181–190 (2010)
10. Xu, Z.S., Yager, R.R.: Intuitionistic and interval-valued intuitionistic fuzzy preference relations and their measures of similarity for the evaluation of agreement within a group. Fuzzy Optim. Decis. Making **8**, 123–139 (2009)
11. Zhang, Q.S., Jiang, S.Y., Jia, B.G., Luo, S.H.: Some information measures for interval-valued intuitionistic fuzzy sets. Inf. Sci. **180**, 5130–5145 (2010)
12. Zhang, Z.H., Wang, M., Hu, Y., Yang, J.Y., Ye, Y.P., Li, Y.F.: A dynamic interval -valued intuitionistic fuzzy sets applied to pattern recognition. Math. Prob. Eng. **2013**(6), no.408012, 1–16 (2013)
13. Zhang, Z.H., Yang, J.Y., Ye, Y.P., Hu, Y., Zhang, Q.S.: Intuitionistic fuzzy sets with double parameters and its application to pattern recognition. Inf. Technol. J. **11**(3), 313–318 (2012)
14. Xu, Z.S.: Dynamic intuitionistic fuzzy multiple attribute decision making. Int. J. Approximate Reasoning **28**, 246–262 (2008)
15. Wei, G.W.: Some geometric aggregation functions and their application to dynamic attribute decision making in intuitionistic fuzzy setting. Int. J. Uncertainty, Fuzziness Knowl. Based Syst. **17**, 251–267 (2009)
16. Su, Z.X., Chen, M.Y., Xia, G.P., Wang, L.: An interactive method for dynamic intuitionistic fuzzy multi-attribute group decision making. Expert Syst. Appl. **38**, 15286–15295 (2011)
17. Zhang, Z.H., Yang, J.Y., Ye, Y.P., Hu, Y., Zhang, Q.S: Intuitionistic fuzzy sets with double parameters and its application to dynamic multiple attribute decision making. Inf. Int. Interdisc. J. **15**(6), 2479–2486 (2012)
18. Zhang, Z.H., Yang, J.Y., Ye, Y.P., Hu, Y., Zhang, Q.S.: A scoring function of intuitionistic fuzzy sets with double parameters and its application to multiple attribute decision making. Inf. Int. Interdisc. J. **15**(11A), 4443–4450 (2012)
19. Hu, Y., Mo, X.Z., Zhang, X.Z., Zeng, Y.R., Du, J.F., Xie, K.: Intelligent analysis model for outsourced software project risk using constraint-based Bayesian network. J. Softw. **7**(2), 440–449 (2012)
20. Hu, Y., Zhang, X.Z., Nagi, E.W.T., Cai, R.C., Liu, M.: Software project risk analysis using Bayesian networks with causality constraints. Decis. Support Syst. **56**(12), 439–449 (2013)
21. Hu, Y., Du, J.F., Zhang, X.Z., Hao, X.L., Nagi, E.W.T., Fan, M., Liu, M.: An integrative framework for intelligent software project risk planning. Decis. Support Syst. **55**(11), 927–937 (2013)
22. Hu, Y., Feng, B., Mo, X.Z., Zhang, X.Z., Nagi, E.W.T., Fan, M., Liu, M.: Cost-Sensitive and ensemble-based prediction model for outsourced software project risk prediction. Decis. Support Syst. **72**(2), 11–23 (2015)
23. Wallace, L., Keil, M., Rai, A.: Understanding software project risk: A cluster analysis. Inf. Manage. **42**(1), 115–125 (2004)
24. Nidumolu, S.: The effect of coordination and uncertainty on software project performance: Residual performance risk as an intervening variable. Inf. Syst. Res. **6**(3), 191–219 (1995)
25. Xia, W.D., Lee, G.: Complexity of information systems development projects: Conceptualization and measurement development. J. Manage. Inf. Syst. **22**(1), 45–84 (2005)
26. Wallace, L., Keil, M.: Software project risks and their effect on outcomes. Commun. ACM **47** (4), 68–73 (2004)
27. Schmidt, R., Lyytinen, K., Keil, M.: Identifying software project risks: An international Delphi study. J. Manage. Inf. Syst. **17**(4), 5–36 (2001)

28. Boehm, B.W.: Software risk management: principles and practices. IEEE Softw. **8**(1), 32–41 (1991)
29. Jiang, J.J., Klein, G.: An exploration of the relationship between software development process maturity and project performance. Inf. Manage. **41**(3), 279–288 (2004)
30. Karolak, D.W.: Software Engineering Risk Management. IEEE Computer Society Press, Los Alamitos (1996)

Weighted Min-Max Programming Subject to Max-Product Fuzzy Relation Equations

Xiao-Peng Yang, Ying-Han Hong, Xue-Gang Zhou and Bing-Yuan Cao

Abstract Weighted min-max programming subject to max-product fuzzy relation equations is investigated in this paper. For solving the proposed problem, we introduce concepts of discrimination matrix and solution matrix and study some of their properties. Based on these matrices, solution method is developed to find the optimal solution of the fuzzy relation weighted min-max programming problem. The solution method is illustrated by a numerical example.

Keywords Weighted min-max programming · Fuzzy relation equation · Max-product composition

1 Introduction

Fuzzy relation equation (FRE) with max-min composition was firstly introduced by Sanchez [1]. The solution set of a consistent system of FREs is determined by one maximum solution and a finite number of minimal solutions. Thus it is usually a non-convex set. Computing the maximum solution is easy, while solving all the minimal solutions is not trivial. Many methods were developed for obtaining the minimal solution set of FREs [2–4]. In some cases, the max-product composition was be superior to the max-min composition [5]. In [6] the authors presented some theoretical results on max-product FRE. Efficient procedure was proposed in [7] for obtaining the minimal solution set to system of max-product FREs. Wu and Guu [8]

X.-P. Yang · Y.-H. Hong
School of Mathematics and Statistics, Hanshan Normal University,
Chaozhou 521041, China

X.-P. Yang · X.-G. Zhou · B.-Y. Cao (✉)
School of Mathematics and Information Science, Guangzhou University,
Guangzhou 510006, China
e-mail: caobingy@163.com

B.-Y. Cao
Guangzhou Vocational College of Science and Technology, Guangzhou 510550, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_28

improved the solution method by applying some rules to reduce the problem. Matrix-pattern-based algorithm, for solving all the minimal solutions to FREs with either max-min or max-product, was investigated and implemented in a computer program by Luoh and Wang [9]. It was shown in [10] computing the minimal solution set of max-product FREs was equivalent to the covering problem.

Minimizing a linear function with max-product FREs constraint was considered by Loetamonphong and Fang [11] for the first time. The main problem was divided into tow subproblems. One of them was handled by the branch-and-bound method and the other one could be easily solved by the maximum solution of the constraint. Qu and Wang [12] studied the same issue by improving the solution method. Besides, optimization problem with nonlinear objective function and max-product FREs or fuzzy relation inequalities (FRIs) was also investigated, e.g. fuzzy relation quadratic programming [13] and geometric programming problems [14]. Recently, Yang et al. [15] introduced the single-variable term semi-latticized geometric programming subject to max-product FREs with application in peer-to-peer network system. Matrix approach was developed for obtaining the optimal solution of the proposed problem. Considering the limitation of total download speed of each user, the peer-to-peer file sharing system could be reduced into a system of fuzzy relation inequalities with addition-min composition [16]. Relevant optimization problems were investigated for the purpose of reducing network congestion in such system [17–19].

In this paper an optimal management model in WiFi terminal system will be considered. Suppose there are n terminals, i.e. $A_1, A_2, \dots, A_n$ in such system. A_j sends out information to other terminals with signal intensity x_j, $j = 1, 2, \dots, n$. When the signal reaches A_i, the intensity becomes $a_{ij}x_j$, where $a_{ij} \in [0, 1]$. In general, A_i will receive the signal with strongest intensity. On the other hand, if the requirement of signal intensity of A_i is b_i, $i = 1, 2, \dots, m$, $m \leq n$, then by normalizing the parameters and variables, the WiFi terminal system can be reduced into the following system of max-product fuzzy relation equations:

$$\begin{cases} a_{11}x_1 \vee a_{12}x_2 \vee \cdots \vee a_{1n}x_n = b_1, \\ a_{21}x_1 \vee a_{22}x_2 \vee \cdots \vee a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 \vee a_{m2}x_2 \vee \cdots \vee a_{mn}x_n = b_m. \end{cases} \tag{1}$$

or

$$A \circ x^T = b^T,$$

where $A = (a_{ij})_{m \times n}$, $x = (x_1, x_2, \dots, x_n)$, $b = (b_1, b_2, \dots, b_m)$, and $a_{ij}, x_j \in [0, 1]$, $b_i \in (0, 1]$, $i \in I, j \in J$. $I = \{1, 2, \dots, m\}$ and $J = \{1, 2, \dots, n\}$ are two index sets. $c_j \in (0, 1]$, $j \in J$. Furthermore, one of the optimal management objective is to minimize the biggest weighted signal intensities, i.e. $c_1x_1 \vee c_2x_2 \vee \cdots \vee c_nx_n$.

We establish the following weighted min-max programming subject to max-product fuzzy relation equations,

$$\begin{aligned} \min\ z(x) &= c_1x_1 \vee c_2x_2 \vee \cdots \vee c_nx_n \\ \text{s.t. } A \circ x^T &= b^T. \end{aligned} \tag{2}$$

where $A \circ x^T = b^T$ is as system (1).

The rest of the paper is organized as follows. Section 2 provides some necessary concepts and results on max-product fuzzy relation equations. Solution method for the proposed problem is developed in Sect. 3. In Sect. 4 we present a numerical example to illustrate our solution method. Simple conclusion lies in Sect. 5.

2 Preliminaries

In this section, we review some relevant concepts and results on max-product fuzzy relation equations.

Definition 1 For $x, y \in R^n$, we write $x \leq y$ $(y \geq x)$, if $x_j \leq y_j$ $(y_j \geq x_j)$ holds for all $j \in \{1, 2, \ldots, n\}$, and $x < y$ $(y > x)$, if $x_j < y_j$ $(y_j > x_j)$ holds for all $j \in \{1, 2, \cdots, n\}$.

It is easy to verify that $\leqslant$ is a partial order relation between R^n. We denote the solution set of (5) by $X(A, b) = \{x \in [0, 1]^n | A \circ x = b\}$.

Definition 2 [20] $A \circ x = b$ is said to be consistent (or inconsistent) if and only if $X(A, b) \neq \emptyset$ (or $X(A, b) = \emptyset$).

Define

$$a_{ij} \circledast^{-1} b_i = \begin{cases} \frac{b_i}{a_{ij}}, & a_{ij} > b_i, \\ 1, & a_{ij} \leq b_i, \end{cases} \tag{3}$$

where $\circledast^{-1}$ is an operator defined on [0,1], $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Let $\hat{x} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)^T$, where

$$\hat{x}_j = \bigwedge_{i=1}^{m} (a_{ij} \circledast^{-1} b_i), \tag{4}$$

$j = 1, 2, \ldots, n$. Then the consistency of $A \circ x = b$ can be verified by the following Proposition 1.

Proposition 1 [20] *$A \circ x = b$ is consistent if and only if $A \circ \hat{x} = b$.*

Remark 1 Considering the practical background, we always assume $b > (0, 0, \ldots, 0)^T$. As a result, it follows that $\hat{x} > (0, 0, \ldots, 0)^T$, i.e. $\hat{x}_j > 0$, $\forall j \in J$, according to (3) and (4).

Definition 3 A solution $\hat{x}$ of $A \circ x = b$ is said to be a maximum solution if and only if $x \leqslant \hat{x}$ for all $x \in X(A, b)$. A solution $\check{x}$ of $A \circ x = b$ is said to be a minimal solution if and only if $x \leqslant \check{x}$ implies $x = \check{x}$ for any $x \in X(A, b)$.

It is well known that $\hat{x}$ is the unique maximum solution of $A \circ x = b$ if $A \circ \hat{x} = b$. The maximum solution $\hat{x}$ may be easily obtained following (6) and (7). If $A \circ x = b$ has a minimal solution, the minimal solutions are usually not unique. If we denote the set of all the minimal solutions of $A \circ x = b$ by $\check{X}(A, b)$, then we have

$$X(A,b) = \bigcup_{\check{x} \in \check{X}(A,b)} \{x | \check{x} \leq x \leq \hat{x}\}. \tag{5}$$

Proposition 2 [11, 20] *If $x \in X(A, b)$, then for each $i \in I$, there exists $j_0 \in J$ such that $a_{ij_0} x_{j_0} = b_i$ and $a_{ij} x_j \leq b_i$ $\forall j \in J$.*

Based on Propositions 1 and 2, we can define the index sets as follows:

$$J_i = \{j \in J | a_{ij} \hat{x}_j = b_i\}, \quad \forall i \in I, \tag{6}$$

and

$$\Lambda = J_1 \times J_2 \times \cdots \times J_m. \tag{7}$$

The consistence of system (2) can be easily checked by the following Corollaries 1 and 2.

Corollary 1 [11] *If $X(A, b) \neq \emptyset$, then $J_i \neq \emptyset, \forall i \in I$.*

Corollary 2 [11] *If $X(A, b) \neq \emptyset$, then $\Lambda \neq \emptyset$.*

3 Solution Method to Problem (2)

Obviously, when system (1) is inconsistent, problem (2) has no optimal solution. In this section, we develop a novel solution method for solving an optimal solution of problem (2) with the constraint being consistent system of max-product fuzzy relation equations.

Let $\hat{x}$ be the maximum solution of consistent system (1) and J_i be the index set obtained by (6), $i \in I$.

Definition 4 (*Discrimination Matrix*) A matrix $D = (d_{ij})_{m \times n}$ is said to be the discrimination matrix of system (1), if

$$d_{ij} = \begin{cases} \hat{x}_j, & \text{if } j \in J_i, \\ 0, & \text{if } j \notin J_i, \end{cases} \qquad \forall i \in I, j \in J. \tag{8}$$

Remark 2 In the discrimination matrix D, it follows from Remark 1 that $d_{ij} > 0$ if and only if $j \in J_i$.

Theorem 1 [20] *System (2) is consistent if and only if each row in the discrimination matrix D has at least one nonzero element.*

Definition 5 (*Solution Matrix*) Let D be the discrimination matrix of system (1) and $S = (s_{ij})_{m\times n}$, where $s_{ij} \in \{0, d_{ij}\}$. S is said to be a solution matrix with respect to D, if for any $i \in I$, there exists an unique $j_i \in J$ such that $s_{ij_i} \neq 0$.

Following Definition 5 it is obvious that each row in the solution matrix S has an unique nonzero element, i.e. $s_{ij_i} = d_{ij_i} > 0, \forall i \in I$. For the discrimination matrix D of system (1), we denote the set of all solution matrices with respect to D by $\mathbb{S}_D$. Let $S = (s_{ij})$ be an arbitrary solution matrix in $\mathbb{S}_D$. We define $x^S = (x_1^S, x_2^S, \ldots, x_n^S)$, where

$$x_j^S = \bigvee_{i\in I} s_{ij}, \quad \forall j \in J. \tag{9}$$

It is shown in the following Theorem 2 that x^S is exactly a solution of system (1). Due to this reason, x^S is said to be the solution corresponding to S.

Theorem 2 *Let $S \in \mathbb{S}_D$ be a solution matrix and x^S be obtained by (9). Then x^S is a solution of system (1).*

Proof Obviously $x^S \in X$. We just need to verify that x^S satisfies the constraint in problem (2). For arbitrary $k \in I$, according to Definition 5, there exists an unique $j_k \in J$, such that

$$s_{kj_k} \neq 0.$$

By Definitions 4 and 5, it follows that

$$s_{kj_k} = d_{kj_k} = \hat{x}_{j_k} > 0, \tag{10}$$

and

$$j_k \in J_k. \tag{11}$$

Again by Definitions 4 and 5, we get

$$s_{ij} \leq d_{ij} \leq \hat{x}_j, \quad \forall j \in J.$$

Thus

$$x_j^S = \bigvee_{i\in I} s_{ij} \leq \bigvee_{i\in I} \hat{x}_j = \hat{x}_j, \quad \forall j \in J. \tag{12}$$

Notice that $0 \leq a_{kj} \leq 1$, for any $j \in J$. Applying (12) we have

$$a_{k1}x_1^S \vee a_{k2}x_2^S \vee \cdots \vee a_{kn}x_n^S \leq a_{k1}\hat{x}_1 \vee a_{k2}\hat{x}_2 \vee \cdots \vee a_{kn}\hat{x}_n. \tag{13}$$

Since $\hat{x}$ is the maximum solution of system (1), it holds that

$$a_{k1}\hat{x}_1 \vee a_{k2}\hat{x}_2 \vee \cdots \vee a_{kn}\hat{x}_n = b_k. \tag{14}$$

Combining (13) and (14), it is immediate that

$$a_{k1}x_1^S \vee a_{k2}x_2^S \vee \cdots \vee a_{kn}x_n^S \leq b_k,$$

i.e.

$$a_{kj}x_j^S \leq b_k, \quad \forall j \in J. \tag{15}$$

On the other hand, it follows from (10) that

$$x_{j_k}^S = \bigvee_{i \in I} s_{ij_k} \geq s_{kj_k} = \hat{x}_{j_k}. \tag{16}$$

Inequalities (12) and (16) contribute to

$$x_{j_k}^S = \hat{x}_{j_k}.$$

Considering (6) and (11), we get

$$a_{kj_k}x_{j_k}^S = a_{kj_k}\hat{x}_{j_k} = b_k. \tag{17}$$

Consequently, according to Proposition 2, it follows from (15) and (17) that x^S is a solution of system (1). □

Theorem 3 *If system (1) is consistent, then its solution set is*

$$X(A, b) = \bigcup_{S \in \mathbb{S}_D} \{x | x^S \leq x \leq \hat{x}\},$$

where D is the discrimination matrix and $\hat{x}$ is the maximum solution of system (1).

Proof It is clear that $\bigcup_{S \in \mathbb{S}_D} \{x | x^S \leq x \leq \hat{x}\} \subseteq X(A, b)$. In order to complete the proof, we need to prove $X(A, b) \subseteq \bigcup_{S \in \mathbb{S}_D} \{x | x^S \leq x \leq \hat{x}\}$.

Let $y \in X(A, b)$ be an arbitrary solution of system (1). Following Proposition 2, for any $i \in I$, there exists $j_i \in J$ such that

$$a_{ij_i}y_{j_i} = b_i. \tag{18}$$

Since $\hat{x}$ is the maximum solution of system (1), we have

$$a_{i1}\hat{x}_1 \vee a_{i2}\hat{x}_2 \vee \cdots \vee a_{in}\hat{x}_n = b_i, \quad \forall i \in I, \tag{19}$$

and

$$\hat{x}_j \geq y_j, \quad \forall j \in J. \tag{20}$$

Equation (19) indicates

$$a_{ij_i}\hat{x}_{j_i} \leq b_i, \quad \forall i \in I,$$

while Equality (18) and Inequality (20) imply that

$$a_{ij_i}\hat{x}_{j_i} \geq a_{ij_i}y_{j_i} = b_i, \quad \forall i \in I.$$

Hence

$$a_{ij_i}\hat{x}_{j_i} = b_i, \quad \forall i \in I. \tag{21}$$

Here, it should be pointed out that $j_i \in J_i$, according to (6). Besides, observing Equalities (18) and (21), it is easy to check that

$$\hat{x}_{j_i} = y_{j_i}, \quad \forall i \in I, \tag{22}$$

since $b_i > 0$.

Next, based on $j_1, j_2, \ldots, j_m$, we construct matrix $S^y = (s_{ij}^y)_{m\times n}$ as follows:

$$s_{ij}^y = \begin{cases} d_{ij}, & \text{if } j = j_i, \\ 0, & \text{if } j \neq j_i, \end{cases} \tag{23}$$

for any $i \in I$. Obviously $j_i \in J_i$ indicates

$$s_{ij_i}^y = d_{ij_i} = \hat{x}_{j_i} > 0, \quad \forall i \in I. \tag{24}$$

On the other hand, for arbitrary $i \in I$, it holds that

$$s_{ij}^y = 0, \quad \text{for any } j \neq j_i. \tag{25}$$

Thus S^y is a solution matrix of system (1), i.e. $S^y \in \mathbb{S}_D$.

Let x^{S^y} be the solution corresponding to S^y. Then we have

(i) when $j \in \{j_1, j_2, \ldots, j_m\}$, there exists some $k \in I$ such that $j_k = j$. Furthermore, it follows from (22) that $\hat{x}_j = y_j$. According to (24),

$$x_j^{S^y} = x_{j_k}^{S^y} = \bigvee_{i\in I} s_{ij_k}^y \geq s_{kj_k}^y = \hat{x}_{j_k} = \hat{x}_j. \tag{26}$$

On the other hand, it is obvious that $x_j^{S^y} \leq \hat{x}_j$ for any $j \in J$, since $\hat{x}$ is the maximum solution. Hence

$$x_j^{S^y} = \hat{x}_j = y_j.$$

(ii) when $j \notin \{j_1, j_2, \dots, j_m\}$, by (25) we get

$$x_j^{S^y} = \bigvee_{i \in I} s_{ij}^y = \bigvee_{i \in I} 0 = 0 \le y_j. \tag{27}$$

Cases (i) and (ii) imply that $x_j^{S^y} \le y_j$ for arbitrary $j \in J$. Consequently,

$$x^{S^y} \le y \le \hat{x},$$

i.e. $y \in \bigcup_{S \in \mathbb{S}_D} \{x | x^S \le x \le \hat{x}\}$, which indicates $X(A, b) \subseteq \bigcup_{S \in \mathbb{S}_D} \{x | x^S \le x \le \hat{x}\}$. □

Since $J_1, J_2, \dots, J_m$ are finite sets, there exists at least one $j_i^* \in J_i$ such that

$$c_{j_i^*} d_{ij_i^*} = \bigwedge_{j \in J_i} \{c_j d_{ij}\}, \tag{28}$$

for any $i \in I$. According to Remark 2, $j_i^* \in J_i$ indicates $d_{ij_i^*} > 0$, $\forall i \in I$. Based on $j_1^*, j_2^*, \dots, j_m^*$, we construct the matrix $S^* = (s_{ij}^*)_{m \times n}$ as follows,

$$s_{ij}^* = \begin{cases} d_{ij}, & \text{if } j = j_i^*, \\ 0, & \text{if } j \ne j_i^*, \end{cases} \tag{29}$$

for any $i \in I$. It is easy to verify that S^* is a solution matrix of system (1), i.e. $S^* \in \mathbb{S}_D$. In the following S^* is said to be an optimal solution matrix.

Theorem 4 *Let S^* be an optimal solution matrix defined by (29) and x^* be the solution corresponding to S^*. Then x^* is an optimal solution of problem (2).*

Proof (Feasibility) As S^* is a solution matrix of system (1), it follows from Theorem 2 that $x^* \in X(A, b)$ is a feasibility solution of problem (2).

(Optimality) For any $l \in I$, we have both

$$x_{j_l^*}^* \le \hat{x}_{j_l^*}$$

and

$$x_{j_l^*}^* = \bigvee_{i \in I} s_{ij_l^*}^* \ge s_{lj_l^*}^* = d_{lj_l^*} = \hat{x}_{j_l^*}.$$

Thus

$$x_{j_l^*}^* = \hat{x}_{j_l^*}, \quad \forall l \in I. \tag{30}$$

Let $y \in X(A, b)$ be an arbitrary feasible solution of problem (2). Now we verify that $c_j x_j^* \le c_1 y_1 \vee c_2 y_2 \vee \cdots \vee c_n y_n = z(y)$ for any $j \in J$ in two cases.

Case 1. If $j \notin \{j_1^*, j_2^*, \ldots, j_m^*\}$, then according to (29) we have

$$c_j x_j^* = c_j \bigvee_{i \in I} s_{ij}^* = c_j \bigvee_{i \in I} 0 = 0 \le z(y). \tag{31}$$

Case 2. If $j \in \{j_1^*, j_2^*, \ldots, j_m^*\}$, there exists $k \in \{1, 2, \ldots, m\}$ such that $j_k^* = j$. If follows from (30) that

$$x_{j_k^*}^* = \hat{x}_{j_k^*}. \tag{32}$$

Moreover, since $j_k^* \in J_k$, following Definition 4 we have

$$d_{kj_k^*} = \hat{x}_{j_k^*}. \tag{33}$$

Combining Equalities (28), (32) and (33), it is obvious that

$$c_j x_j^* = c_{j_k^*} x_{j_k^*}^* = c_{j_k^*} \hat{x}_{j_k^*} = c_{j_k^*} d_{kj_k^*} = \bigwedge_{j \in J_k} \{c_j d_{kj}\}. \tag{34}$$

Since $y \in X(A, b)$ is a solution of system (1), according to Theorem 3, there exists a solution matrix $S^y = (s_{ij}^y)_{m \times n}$, such that $x^{S^y} \le y \le \hat{x}$, where x^{S^y} is the solution corresponding to S^y. Following Definition 5, for $i = k$, there exists an unique j_k such that $s_{kj_k}^y \ne 0$, which indicates

$$j_k \in J_k \text{ and } s_{kj_k}^y = d_{kj_k} = \hat{x}_{j_k}.$$

Hence

$$\bigwedge_{j \in J_k} \{c_j d_{kj}\} \le c_{j_k} d_{kj_k} = c_{j_k} \hat{x}_{j_k} \le c_{j_k} y_{j_k} \le \bigvee_{j \in J} c_j y_j = z(y). \tag{35}$$

Combining (34) and (35) we get $c_j x_j^* = \bigwedge_{j \in J_k} \{c_j d_{kj}\} \le z(y)$.

Cases 1 and 2 show that $c_j x_j^* \le z(y)$ for any $j \in J$. Consequently,

$$z(x^*) = \bigvee_{j \in J} c_j x_j^* \le \bigvee_{j \in J} z(y) = z(y).$$

This contributes to the optimality of x^* and the proof is completed. □

4 Illustrative Example

Example 1 Consider the following weighted min-max programming subject to max-product fuzzy relation equations:

$$\begin{aligned} &\min z(x) = 0.8x_1 \vee 0.9x_2 \vee 0.5x_3 \vee 0.3x_4 \vee 0.7x_5 \vee 0.6x_6 \vee 0.7x_7 \vee 0.7x_8 \\ &\text{ s.t. } A \circ x^T = b^T, \end{aligned} \tag{36}$$

where $x = (x_1, x_2, \dots, x_8)$, $b = (b_i) = (0.72, 0.64, 0.56, 0.42, 0.42, 0.4)$, and

$$A = (a_{ij}) = \begin{bmatrix} 0.9 & 0.9 & 0.8 & 0.2 & 0.8 & 0.6 & 0.1 & 0.4 \\ 0.5 & 0.8 & 0.7 & 0.4 & 0.7 & 0.8 & 0.3 & 0.8 \\ 0.8 & 0.6 & 0.2 & 0.4 & 0.2 & 0.7 & 0.7 & 0.5 \\ 0.6 & 0.3 & 0.7 & 0.6 & 0.1 & 0.3 & 0.5 & 0.3 \\ 0.6 & 0.2 & 0.5 & 0.5 & 0.1 & 0.4 & 0.7 & 0.2 \\ 0.2 & 0.4 & 0.5 & 0.1 & 0.3 & 0.5 & 0.8 & 0.4 \end{bmatrix}.$$

Solution

By (3) and (4) we get $\hat{x} = (0.7, 0.8, 0.6, 0.7, 0.9, 0.8, 0.5, 0.8)$. It is easy to check that $A \circ \hat{x} = b^T$. Hence system (36) is consistent according to Proposition 1. By (6), the index sets are $J_1 = \{1, 6\}$, $J_2 = \{1, 3, 4\}$, $J_3 = \{2, 6, 8\}$, $J_4 = \{6, 7\}$, $J_5 = \{1\}$, $J_6 = \{2, 5\}$. Based on the maximum solution $\hat{x}$ and $J_1, J_2, \dots, J_6$, the discrimination matrix can be written as

$$D = \begin{bmatrix} 0.7 & 0 & 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0.7 & 0 & 0.6 & 0.7 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 & 0.8 & 0 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 0.8 & 0.5 & 0 \\ 0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0.9 & 0 & 0 & 0 \end{bmatrix}.$$

Now we compute j_i^* $(i = 1, 2, \dots, 6)$ by (28) as follows.

$$\begin{aligned} \bigwedge_{j \in J_1} c_j d_{1j} &= 0.8 \cdot 0.7 \wedge 0.6 \cdot 0.8 = 0.48 = c_6 d_{16}, \\ \bigwedge_{j \in J_2} c_j d_{2j} &= 0.8 \cdot 0.7 \wedge 0.5 \cdot 0.6 \wedge 0.3 \cdot 0.7 = 0.21 = c_4 d_{24}, \\ \bigwedge_{j \in J_3} c_j d_{3j} &= 0.9 \cdot 0.8 \wedge 0.6 \cdot 0.8 \wedge 0.7 \cdot 0.8 = 0.48 = c_6 d_{36}, \\ \bigwedge_{j \in J_4} c_j d_{4j} &= 0.6 \cdot 0.5 \wedge 0.7 \cdot 0.8 = 0.30 = c_6 d_{46}, \\ \bigwedge_{j \in J_5} c_j d_{5j} &= 0.8 \cdot 0.7 = 0.56 = c_1 d_{51}, \\ \bigwedge_{j \in J_6} c_j d_{6j} &= 0.9 \cdot 0.8 \wedge 0.7 \cdot 0.9 = 0.63 = c_5 d_{65}. \end{aligned} \tag{37}$$

That is, $j_1^* = 6, j_2^* = 4, j_3^* = 6, j_4^* = 6, j_5^* = 1, j_6^* = 5$. Based on $j_1^*, j_2^*, \dots, j_6^*$ and the discrimination matrix D, the optimal solution matrix may be established as follows

$$S^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.9 & 0 & 0 & 0 \end{bmatrix}.$$

The solution corresponding to S^* is $x^* = (0.7, 0, 0, 0.7, 0.9, 0.8, 0, 0)$. The vector x^* is exactly an optimal solution of problem (36). That is, the normalized signal intensities are $0.7, 0, 0, 0.7, 0.9, 0.8, 0, 0$ respectively.

5 Conclusion

Based on the discrimination matrix and solution of a system of max-product fuzzy relation equations, we develop a solution method to the corresponding weighted min-max programming problem. As shown in Sect. 3, for each solution matrix S, there exists a solution (of system (1)) corresponding to S. Moreover, the solution corresponding to the optimal solution matrix S^* is exactly an optimal solution of problem (2).

Acknowledgments Thanks to the support by the Innovation and Building Strong School Project of Colleges of Guangdong Province (2015KQNCX094).

References

1. Sanchez, E.: Resolution of composite fuzzy relation equations. Inf. Control **30**, 38–48 (1976)
2. Shieh, B.-S.: Deriving minimal solutions for fuzzy relation equations with max-product composition. Inf. Sci. **178**, 3766–3774 (2008)
3. Matusiewicz, Z., Drewniak, J.: Increasing continuous operations in fuzzy max-* equations and inequalieties. Fuzzy Sets Syst. **232**, 120–133 (2013)
4. Peeva, K.: Resolution of fuzzy relational equations-method, algorithm and software with applications. Inf. Sci. **234**, 44–63 (2013)
5. Zimmermann, H.-J.: Fuzzy Set Theory and Its Applications. Kluwer Academic Publishers, Boston (1991)
6. Bourke, M.M., Grant Fisher, D.: Solution algorithms for fuzzy relational equations with max-product composition Fuzzy Sets Syst. **94**, 61–69 (1998)
7. Loetamonphong, J., Fang, S.-C.: An efficient solution procedure for fuzzy relational equations with max-product composition. IEEE Trans. Fuzzy Syst. **7**, 441–445 (1999)
8. Wu, Y.-K., Guu, S.-M.: Finding the complete set of minimal solution for fuzzy relational equations with max-product compostion. Int. J. Oper. Res. **1**(1), 29–36 (2004)
9. Luoh, L., Wang, W.-J.: Matrix-pattern-based computer algorithm for solving fuzzy relation equations. IEEE Trans. Fuzzy Syst. **11**(1), 100–108 (2003)
10. Markovskii, A.V.: On the relation between equations with max-product composition and the covering problem. Fuzzy Sets Syst. **153**, 261–273 (2005)

11. Loetamonphong, J., Fang, S.-C.: Optimization of fuzzy relation equations with max-product composition. Fuzzy Sets Syst. **118**, 509–517 (2001)
12. Qu, X.-B., Wang, X.-P.: Minimization of linear objective functions under the constraints expressed by a system of fuzzy relation equations. Inf. Sci. **178**, 3482–3490 (2008)
13. Abbasi Molai, A.: A new algorithm for resolution of the quadratic programming problem with fuzzy relation inequality constraints. Comput. Ind. Eng. **72**, 306–314 (2014)
14. Shivanian, E., Khorram, E.: Monomial geometric programming with fuzzy relation inequality constraints with max-product composition. Comput. Ind. Eng. **56**, 1386–1392 (2009)
15. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.: Single-variable term semi-latticized fuzzy relation geometric programming with max-product operator. Inf. Sci. **325**, 271–287 (2015)
16. Li, J.-X., Yang, S.-J.: Fuzzy relation equalities about the data transmission mechanism in bittorrent-like peer-to-peer file sharing systems. In: Proceedings of the 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2012, pp. 452–456
17. Yang, S.-J.: An algorithm for minimizing a linear objective function subject to the fuzzy relation inequalities with addition-min composition. Fuzzy Sets Syst. **255**, 41–51 (2014)
18. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.: Multi-level linear programming subject to addition-min fuzzy relation inequalities with application in Peer-to-Peer file sharing system. J. Intell. Fuzzy Syst. **28**, 2679–2689 (2015)
19. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.: Min-max programming problem subject to addition-min fuzzy relation inequalities. IEEE Trans. Fuzzy Syst. **24**(1), 111–119 (2016)
20. Cao, B.-Y.: Optimal Models and Methods with Fuzzy Quantities. Springer, Berlin (2010)

Part III
Internet and Big Data Applications

A Forecasting Approach of Fuzzy Time Series Model Based on a New Data Fuzzification

Gang Chen, Li-hong Yang and Xue Yang

Abstract In view of the research about fuzzy time series models, the existing models are lack of objective data fuzzification and sensitivity. In this paper, firstly, a new method of defining fuzzy sets present is set up and six new fuzzy sets are given. Secondly, the rules of data fuzzification are defined. Finally, the model is used to forecast the enrollments of the University of Alabama. It is shown that the proposed model gets a higher forecasting accuracy than those which use traditional methods to forecast.

Keywords Fuzzy time series · Fuzzy sets definition · Data fuzzification · Efficiency

1 Introduction

Traditional forecasting methods can deal with many forecasting cases, but they cannot solve forecasting problems in which the historical data are linguistic values. Song and Chrisom [1–3], they first presented the concepts of fuzzy time series based on the fuzzy set theory [4]. Song and Chrisom explored forecasting of fuzzy time series with enrollment data of the University of Alabama and proposed a forecasting framework. Since the work of Song and Chrisom, numerous studies have been conducted to improve forecasting accuracy or reduce computational overhead.

Huarng [5], Huarng and Yu [6], Cheng, Chang, and Yeh [7], Chen and Hsu [8], Chen and Chung [9, 10], Li, Cheng, and Lin [11] focused on the partition of

G. Chen (✉) · L. Yang · X. Yang
Department of Mathematics, Dalian Maritime University,
Dalian 116026, China
e-mail: chengang@dlmu.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_29

discourse and the interval length to improve forecasting accuracy. Huarng and Yu presented ratio-based lengths of intervals instead of equal lengths of intervals to improve fuzzy time series forecasting [10], Cheng, Chang and Yeh proposed Minimize Entropy Principle Approach (MEPA) and Trapezoid Fuzzification Approach (TFA) to partition the universe of discourse and build membership functions [7]. Chen and Hsu used a two phase partitioning method based on the statistical distributions of the historical data. It usually resulted in numerous intervals and complicated the task of defuzzification [8]. Chen and Chung used genetic algorithms to adjust the length of each interval for first-order and high-order forecasting models, respectively [9, 10]. Li, Cheng and Lin applied fuzzy c-means clustering to deal with interval partitioning [11].

From the above studies, we can see that there has been no change in fuzzy sets definition and data fuzzification, the method of subjective definition of fuzzy sets has some inevitable questions: In process of data fuzzification, different data will be fuzzified into the same fuzzy set, resulting in the data passivation. This is not well when reflect the sensitivity of the data. Therefore, it is especially important to define the fuzzy sets and fuzzify the data. In this paper, we set up a general rule of fuzzy sets in fuzzy time series and give six new fuzzy sets. Further, we come into being a new method of data fuzzification. The new method can make fuzzy sets change with the data change, resulting in enhance sensitivity and interpretability of the data.

2 Related Literatures

In this section, we briefly review some concepts of fuzzy time series from Song and Chrisom [1].

Definition 1 $Y(t)(t = \ldots, 0, 1, 2, \ldots)$, a subset of R. Let $Y(t)$ be the universe of discourse defined by the fuzzy set $f_i(t)$. If $F(t)$ consists of $f_i(t)(i = 1, 2, \ldots)$, $F(t)$ is defined as a fuzzy time series on $Y(t)(t = \ldots, 0, 1, 2, \ldots)$.

Definition 2 If there exists a fuzzy relationship $R(t-1, t)$, such that $F(t) = F(t-1) \times R(t-1, t)$ where $\times$ is an operator, then $F(t)$ is said to be caused by $F(t-1)$. The relationship between $F(t)$ and $F(t-1)$ can be denoted by $F(t-1) \rightarrow F(t)$.

Definition 3 Suppose $F(t-1) = A_i$ and $F(t) = A_j$ a fuzzy logical relationship is defined as $A_i \rightarrow A_j$, where A_i is named as left-hand side of the fuzzy logical relationship and A_j is the right-hand side. Note the repeated fuzzy logical relationships are removed.

3 A New Method of Data Fuzzification

The study of fuzzy time series has increasingly attracted much attention due to its salient capabilities of tackling uncertainty and vagueness inherent in the data collected. However, when using fuzzy time series to forecast, it is necessary to do the data fuzzification.

The fuzzy sets definition and data fuzzification in many articles are given by Song [2]. However, if little difference among the sample data, it will make some of the data fall on the same interval. As the result, the different data is fuzzified into the same fuzzy set. But the sample points in the universe of discourse of distribution are uneven, they have complex internal structure, obviously, this will not be able to better reflect the sensitivity and accuracy of the data. To solve the problem, we set up a general rule of fuzzy sets in fuzzy time series and give six new fuzzy sets.

Definition 4 Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_k\}$ $u_k(x)$ is the membership function, $\forall x_1, x_2 \in u_i,\ 1 \le i \le k$ where $i,\ k \in N^+$ if $x_1 > x_2$, then

$$\begin{cases} u_k(x_2) < u_k(x_1) < 1 & k < i \\ u_k(x_1) = u_k(x_2) = 1 & k = i \\ u_k(x_1) < u_k(x_2) < 1 & k > i \end{cases} \tag{1}$$

A fuzzy set A_k of U is defined by

$$A_k = (u_1(x_k), u_2(x_k), \ldots, u_i(x_k)) \tag{2}$$

where x_k one of the sample data points is, $u_i(x_k)$ is the degree of belongingness of x_k to u_i.

Here, we give six different fuzzy sets which satisfy the Definition 4.

Definition 5 Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_k\}$ $u_k(x)$ is the membership function, $\forall x_1, x_2 \in u_i, 1 \le i \le k$, fuzzy sets which satisfy the Definition 4:

(1) Trapezoidal distribution:

$$u_i(x) = \begin{cases} \frac{d_{i+1} - d_i}{d_{i+1} - x} & x < d_i \\ 1 & d_i \le x \le d_{i+1} \\ \frac{d_{i+1} - d_i}{x - d_i} & x > d_{i+1} \end{cases} \tag{3}$$

(2) K times parabolic type distribution:

$$u_i(x) = \begin{cases} \left(\frac{d_{i+1} - d_i}{d_{i+1} - x}\right)^k & x < d_i \\ 1 & d_i \le x \le d_{i+1} \\ \left(\frac{d_{i+1} - d_i}{x - d_i}\right)^k & x > d_{i+1} \end{cases} \tag{4}$$

(3) Hyperbolic distribution:

$$u_i(x) = \frac{d_{i+1} - d_i}{|x - d_i| + |d_{i+1} - x|} \tag{5}$$

(4) Cauchy distribution:

$$u_i(x) = \begin{cases} \frac{1}{1+\alpha(d_i - x)^\beta} & x < d_i \\ 1 & d_i \le x \le d_{i+1} \\ \frac{1}{1+\alpha(x - d_{i+1})^\beta} & x > d_{i+1} \end{cases} \tag{6}$$

(5) Gaussian distribution:

$$u_i(x) = \begin{cases} e^{-(\frac{d_i - x}{\sigma})^2} & x < d_i \\ 1 & d_i \le x \le d_{i+1} \\ e^{-(\frac{x - d_{i+1}}{\sigma})^2} & x > d_{i+1} \end{cases} \tag{7}$$

(6) Ridge Type distribution:

$$u_i(x) = \begin{cases} \frac{1}{2}\sin\frac{\pi(d_{i+1} - d_i)}{2(d_{i+1} - x)} & x < d_i \\ 1 & d_i \le x \le d_{i+1} \\ \frac{1}{2}\sin\frac{\pi(d_{i+1} - d_i)}{2(x - d_i)} & x > d_{i+1} \end{cases} \tag{8}$$

where $k > 0, \alpha > 0, \beta > 0\, d_i$ and $d_{i+1}(i \in N^+)$ are the boundary points of intervals.

The data fuzzifying rules are as follows: if there is a data point x_k, and x_k belongs to interval u_i, the membership values of the data point to interval u_i will be 1, and to other intervals are less than 1, namely, $u_i(x_k) = 1,\ u_j(x_k) < 1,\ i,j = 1,2\ldots,c,\ i \ne j$, so the data point is fuzzified into $A_k, A_k = (u_1(x_k), u_2(x_k), \ldots, u_c(x_k))$. Finally, according to above rules to fuzzify all the sample data points, we get the corresponding fuzzy time series as follows:

$$\begin{cases} A_1 = (u_1(x_1), u_2(x_1), \ldots u_c(x_1)) \\ A_2 = (u_1(x_2), u_2(x_2), \ldots u_c(x_2)) \\ \cdots \\ A_n = (u_1(x_n), u_2(x_n), \ldots u_c(x_n)), \end{cases} \tag{9}$$

where $x_1, x_2, \ldots x_n$ are the sample data points.

By new fuzzy sets and data fuzzification method with distance, we can see that the fuzzy sets are different and sensitive to the data's variation. As results, this method can objectively better reflect the real data structure distribution; what's more, enhance the sensitivity and interpretability of the data.

4 The New Method of Fuzzy Time Series Model

The proposed algorithm includes the following four steps to handle forecasting problem:

Step 1 Define the universe of discourse and partition the universe of discourse into several equal length intervals. We define the universe of discourse $U = [D_{\min} - D_1, D_{\max} - D_2]$, where $D_{\min}$ and $D_{\max}$ are the minimum value and the maximum value of the historical training data, respectively, and D_1 and D_2 are proper positive real numbers to divide the universe of discourse U into n equal length intervals $u_1, u_2, \ldots u_c$, and the centers of these intervals are $c_1, c_2, \ldots c_n$, respectively.

Step 2 Define the fuzzy sets and fuzzify the data. According to the above new fuzzy sets definition and the data fuzzification rules, we define fuzzy sets $A_k = (u_1(x_k), u_2(x_k), \ldots, u_c(x_k))$, and get the corresponding fuzzy time series.

Step 3 Build fuzzy logical relationships. Suppose we get a window basis to w years to forecast, then the variation of last year is used to be a criterion and the other variations of w past years are used to form a matrix which is called the operation matrix.
The criterion matrix $C(t)$ and the operation matrix $O^w(t)$ at year t are expressed as follows:

$$C(t) = f(t-1) = [\,C_1 \quad C_2 \quad \ldots \quad C_c\,] \tag{10}$$

$$O^w(t) = \begin{bmatrix} f(t-2) \\ f(t-3) \\ \ldots\ldots \\ f(t-w-1) \end{bmatrix} = \begin{bmatrix} O_{11} & O_{12} & \ldots & O_{1c} \\ O_{21} & O_{22} & \ldots & O_{2c} \\ \ldots & \ldots & \ldots & \ldots \\ O_{(w-1)1} & O_{(w-1)2} & \ldots & O_{(w-1)c} \end{bmatrix} \tag{11}$$

where $0 \le C_j \le 1$, and we can get a relation matrix $R(t)$ by performing $R(t) = O^w(t) \otimes C(t)$, where

$$R(t) = C(t) \otimes O^w(t) = \begin{bmatrix} R_{11} & R_{12} & \ldots & R_{1c} \\ R_{21} & R_{22} & \ldots & R_{2c} \\ \ldots & \ldots & \ldots & \ldots \\ R_{(w-1)1} & R_{(w-1)2} & \ldots & R_{(w-1)c} \end{bmatrix} \tag{12}$$

where $0 \le O_{ij} \le 1,\ 1 \le i \le w-1,\ 1 \le j \le c,\ 1 \le C_c \le 1,\ R_{ij} = O_{ij} \times C_j$, and "×" is the multiplication operation. From the relation matrix $R(t)$, we can know the degree of relationships between last year and the other past years in data variations, then, we can get the forecasting variation of the enrollment of year t, where

$$f(t) = [Max(R_{11}, R_{21}, \ldots, R_{(w-1)1}),\ Max(R_{12}, R_{22}, R_{(w-1)2}), \ldots, Max(R_{1c}, R_{2c}, \ldots, R_{(w-1)c})] \\ = [f_{t1}, f_{t2}, \ldots, f_{tc}] \quad (13)$$

Step 4 Forecast and defuzzify the forecasting outputs. We utilize a more advanced formula, weighted average method to calculate the defuzzified variation F_t:

$$F_t = \sum_{i=1}^{c} c_i f_{ti} / \sum_{i=1}^{c} f_{ti} \quad (14)$$

The forecasted value $X(t)$ at time t is computed as follows:

$$X(t) = X(t-1) + F_t \quad (15)$$

where c_i is the center of the interval u_i, f_{ti}, is the membership value, F_t is forecasting value of the variation, c is the number of the intervals, the actual enrollment of t year is $F(t)$ and the actual enrollment of $t-1$ year is $F(t-1)$.

5 Validation of New Method

The data of historical enrollment of the University of from 1976 year to 1992 year are adopted to illustrate the forecasting process of the new method. We can obtain the variations of the enrollments between any two continuous years as shown in Table 1.

Table 1 Actual enrollments and variations of historical data

Year	Actual enrollments	Variations
1976	15311	+292
1977	15603	+258
1978	15861	+946
1979	16807	+112
1980	16919	−531
1981	16388	−955
1982	15433	+64
1983	15497	−352
1984	15145	+18
1985	15163	+821
1986	15984	+875
1987	16859	+1291
1988	18150	+820
1989	18970	+358
1990	19328	+9
1991	19337	−461
1992	18876	

In order to compare new method with previous other methods, we partition the universe of discourse U into six intervals, and define the fuzzy sets with hyperbolic distribution.

The results of the forecasted enrollments of the University of Alabama are shown in Table 2.

The curve of the actual enrollments and the forecasted enrollments are shown in Fig. 1. The results demonstrate that proposed method generally outperforms the others.

To illustrate the effectiveness of the fuzzy sets defined in this paper. Table 3 shows the forecasted results of our method and Chen et al. [12]. As can be seen from Table 3, the forecasted results of the fuzzy sets defined with Cauchy

Table 2 Actual and forecasted enrollments and errors

Years	Actual enrollments	Forecasted enrollments	Errors (%)
1977	15603	15484	0.77
1978	15861	15893	0.20
1979	16807	16226	3.46
1980	16919	17447	3.12
1981	16388	16311	0.47
1982	15433	16267	5.41
1983	15497	15466	0.20
1984	15145	15655	3.37
1985	15163	15265	0.68
1986	15984	15297	4.30
1987	16859	16635	1.33
1988	18150	17460	3.80
1989	18970	18932	0.20
1990	19328	19545	1.12
1991	19337	19747	2.12
1992	18876	19466	3.12

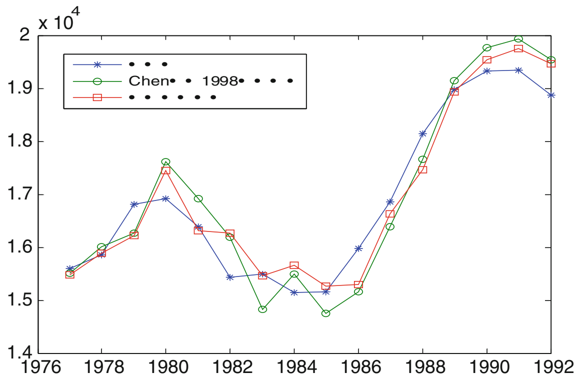


Fig. 1 The curve of the actual and forecasted enrollments

Table 3 Forecast results with different fuzzy sets

Year	Actual enrollments	Chen (1998) [13]	Cauchy distribution: α = 0.5 β = 2	Hyperbolic distribution
1977	15603	15511	15357	15484
1978	15861	16003	16003	15893
1979	16807	16261	16261	16226
1980	16919	17607	17607	17447
1981	16388	16919	16263	16311
1982	15433	16188	16003	16267
1983	15497	14833	15287	15466
1984	15145	15497	15497	15655
1985	15163	14745	15220	15265
1986	15984	15163	15163	15297
1987	16859	16384	16784	16635
1988	18150	17659	17659	17460
1989	18970	19150	19217	18932
1990	19328	19770	19770	19545
1991	19337	19928	19728	19747
1992	18876	19537	19337	19466
MSE		284140	181680	198100

K times parabolic type distribution k = 10	Gaussian distribution	Trapezoidal distribution	Ridge type distribution
15318	15448	15494	15746
15959	15867	15880	15921
16252	16180	16208	16224
17609	17377	17368	17368
16252	16273	16307	16285
15991	16257	16330	16419
14965	15227	15483	15461
15517	15646	15671	15678
15135	15192	15274	15254
15171	15289	15316	15339
16779	16536	16567	16575
17658	17399	17388	17406
19297	18874	18833	18831
19765	19486	19476	19505
19730	19691	19721	19717
19343	19459	19486	19513
200540	205910	209040	218300

distribution are more efficient than other methods, the forecasted results of the fuzzy sets defined with the Gaussian, trapezoidal, Ridge distribution are also superior to the Chen method [12].

Facing different sample data, there are different characters of distribution. We can find a more suitable fuzzy set to reflect the characteristics of the distribution of the sample data. And then, we can obtain a higher forecasting accuracy rate.

6 Conclusion

In this paper, we propose a new method of data fuzzification about fuzzy time series. we set up a general rule of fuzzy sets in fuzzy time series and give six new fuzzy sets. Further, we come into being a new method of data fuzzification. The new method can make fuzzy sets change with the data change, resulting in enhancing sensitivity and interpretability of the data. Thus, we can solve the phenomenon of data passivation generated by the objective definition fuzzy sets. By comparing with Chen's method, the results indicate that our method is more efficient than Chen's method.

References

1. Song, Q., Chissom, B.S.: Fuzzy time series and its models [J]. Fuzzy Sets Syst. **54**(1), 269–277 (1993)
2. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series—Part I [J]. Fuzzy Sets Syst. **54**(1), 1–10 (1993)
3. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series—Part II [J]. Fuzzy Sets Syst. **62**(l), 1–8 (1994)
4. Zadeh, L.: A. Fuzzy sets [J]. Inf. Control **8**, 338–353 (1965)
5. Huarng, K., Yu, T.H.-K.: Ratio-based lengths of intervals to improve fuzzy time series forecasting [J]. IEEE Trans. Syst. Man Cybern. Part B: Cybern. **36**(2), 328–340 (2006)
6. Li, S.-T., Cheng, Y.-C., Lin, S.-Y.: A FCM-based deterministic forecasting model for fuzzy time series [J]. Comput. Math. Appl. **56,** 3052–3063 (2008)
7. Chen, S.M., Chung, N.Y.: Forecasting enrollments of students by using fuzzy time series and genetic algorithms [J]. Inf. Manage. Sci. **17**(3), 1–17 (2006)
8. Optimization [J]. Expert Syst. Appl. **36**(3), 6108–6117 (2009)
9. Chen, S.M., Hsu, C.C.: A new method to forecast enrollments using fuzzy time series [J]. Int. J. Appl. Sci. Eng. **3**, 234–244 (2004)
10. Kuo, I.H., Horng, S.J., Kao, T.W., Lin, T.L., Lee, C.L., Pan, Y.: An improved method for forecasting enrollments based on fuzzy time series and particle swarm
11. Cheng, C.H., Chang, R.J., Yeh, C.A.: Entropy-based and trapezoidal Fuzzification based fuzzy time series approach for forecasting IT project cost [J]. Technol. Forecast. Soc. Change **73**, 524–554 (2006)
12. Singh, P., Borah, B.: An efficient time series forecasting model based on fuzzy time series [J]. Eng. Appl. Artif. Intell. **26**, 2443–2457 (2013)
13. Chen, S.M.: Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Syst. **81**, 311–319 (1996)

Intelligent Control Strategy of Incineration Process Pollution in Municipal Solid Waste

Qian Wu and Hu Xu

Abstract The waste incineration can realize the harmlessness, reduction and resourcfulization of municipal solid waste, and it is not only conducive to the protection of the ecological environment, but also it can obtain better social and economic benefits. Aimed at the pollution resulted in waste incineration, the paper explored the intelligent control strategy in actualizing optimization incineration. In the paper, it discussed the control condition of realizing stable combustion and the evaluation index of incinerating effect, and researched on control strategy, control model and control algorithm. Taking waste incineration process as an example, the simulation experiment demonstrated better control quality of intelligent control strategy. The engineering application and the simulation results show that the presented control strategy is feasible and effective for actualizing optimization control of incineration process in municipal solid waste.

Keywords Municipal solid waste · Stable incineration · Intelligent control strategy

1 Introduction

With the rapid development of economy and city urbanization, it has produced a sharp increase in the amount of the municipal solid waste. The per capita annual yield of municipal solid waste is about 450–500 kg, and its growth rate is more than

Q. Wu (✉)
School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China
e-mail: wuqian80@163.com

H. Xu
College of Automation, Chongqing University, Chongqing 400044, China
e-mail: 1597725908@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_30

8–10 %. Therefore, how to dispose of municipal solid waste is facing an important research topic. Generally, there are four sorts of disposal methods such as general compost, incineration, landfill and comprehensive utilization, and in which, the power generation of municipal solid waste incineration undoubtedly is more convenient to achieve the target of harmlessness, reduction and resourcfulization. But it must make the strict control for the atmospheric pollution from incineration process so as to facilitate the protection of the ecological environment [1, 2]. In the paper, it presented the intelligent control strategy that it can reduce atmospheric pollution from dealing with municipal solid waste incineration, and explored the related control puzzles.

2 Incinerating Puzzle and Its Evaluation Indexes

2.1 Existent Control Puzzles of the Incineration Process

When the municipal solid waste will be as the power combustion raw material, the heat value is very unstable. The incinerator will also change in thermodynamic characteristics because of equipment overhaul, reformation and long-term operation, and it appears probably the puzzles such as difficult ignition, incomplete combustion, firepot scorification, corrosion and secondary pollution aggravation etc. It will lead to be poor in effect of waste incineration, and difficult to control for the pollutants generated by combustion. The fundamental reason is that the cybernetics characteristic of waste incineration process is difficult to accurately grasp, and conventional PID control can not achieve the better control quality. The PID control needs to accurately describe the process characteristics, and establish accurate mathematical model of the process and perfect control loop. Because the combustion process is a very complex physical and chemical process, and also a multi-input multi- output nonlinear process with strong coupling, and it is difficult to establish an accurate model for PID combustion process control. In addition, the stability of incinerator safe operation and combustion process are closely related, if the combustion stability decreased then there will be the combustion efficiency decrease, incineration pollutant emissions increase, secondary pollution and high temperature corrosion aggravating, and even it probably results in serious impact on the safety and economy [3, 4]. Therefore the relevant environmental protection laws and regulations make a series of strict index rules for the waste incineration temperature, content of pollutants in flue gas emissions and thermal burning decay rate of the slag in waste incinerator operation. Focused on achieving the stable combustion, the study on control strategy and control method of effective incineration is of great practical significance for improving the safety and economy.

2.2 Incineration Evaluating Index

In order to evaluate the incineration effect, a series of quantitative evaluation index must be developed so as to examine whether up to the standard after incineration of municipal solid waste in harmlessness, reduction and resourcfulization, and by means of quantitative index, it makes the measure and analysis on incineration treatment effect. The main technical limited-indexes of direct evaluation include the reduction ratio, thermal ignition loss, incineration efficiency, destruction removal efficiency and flue gas emission concentration.

① Reduction ratio
Reduction ratio is defined by expression (1) that the reduced quality after incineration of combustible waste takes the percentage of the total quality in combustible material waste of added waste. This standard is used to measure the reduction effect of incineration waste.

$$MRC = \frac{m_b - m_a}{m_b - m_c} \times 100\,\% \tag{1}$$

In which, MRC is a reduction ratio (%), ma is the quality of incineration residue (kg), mb is the added waste quality (kg), mc is the quality of combustible materials in residue (kg).

② Thermal ignition loss
The thermal ignition loss means that the reduced quality of the incineration residue through 3 h after burning at 600 ± 25 °C takes the percentage of the original quality of incineration residue, and it is computed by expression (2).

$$Q_R = \frac{m_a - m_d}{m_a} \times 100\,\% \tag{2}$$

In which, Q_R is the thermal ignition loss (%), ma is the incineration residue quality at room temperature (kg), md is the incineration residue quality through 3 h after burning at 600 ± 25 °C and cooled to room temperature (kg). It shows that under the condition of high temperature and excess air, when the waste incineration slag is fully oxidized, the incineration slag of unit mass is the amount of being reduced. The greater the thermal ignition loss is, the more incomplete the combustion reaction is, and the poor the incineration effect is. And vice versa, the incineration effect would be better.

③ Incineration efficiency and destruction removal efficiency
In municipal solid waste incineration, it takes the combustion efficiency (CE) as evaluation index that can judge whether to achieve the desired treatment requirements, and it can be shown by expression (3).

$$CE = \frac{[CO_2]}{[CO_2] + [CO]} \times 100\,\% \tag{3}$$

In which, [CO] and [CO_2] are respectively the corresponding value of the gas concentration in the flue gas outtake.

④ Limited index of flue gas emission concentration
In the incineration process of municipal solid waste, it will produce a series of new pollutants, and probably it causes the secondary pollution. For control project of atmosphere pollutant emissions in incineration facilities, it roughly consists of the following four aspects. Smoke dust, it always takes the particles, density, and total carbon content as the control index. Noxious gas, it includes SO_2, HCL, HF, CO and NO_X. Heavy metal element or its compounds, it includes Hg, Cd, Pb, Ni, Cr, As and so on. Organic pollutant, it includes Dioxin and so on.

3 Research of Incineration Stability Control

Through the research of above evaluation indexes, it can be seen that the stable combustion is very important. Stable combustion means that under the condition of being low in waste calorific value and changeable in waste component, it can maintain the necessary combustion temperature so as to ensure that the waste can be fully burnt up. It makes that the heat energy can be utilized effectively and bring secondary pollution minimum, and at the same time, it can reduce the cost of processing equipment operation. The main factors affecting the stable combustion includes the waste characteristic, residence time, turbulence, excess air coefficient etc., in which, the residence time, temperature and turbulence are the main index to reflect the combustion performance, and at the same time, it is also an important index of incinerator control. From analysis of the combustion process characteristics, it can be seen that the incinerator temperature is the most important factor. The high incinerator temperature is conducive to be fully quick dry in waste so as to ensure sufficient pyrolysis process, and promotes the residual carbon burnout to improve waste combustion rate so as to meet the requirement of thermal ignition loss. And high incinerator temperature is also conducive to reduce the harmful pollutants such as dioxin emissions. For example, the temperature range from 300 to 750 °C is extremely easy to generate dioxins, and when the temperature reached 850 °C, after exhaust gas staying 2S, the dioxin about 99.9 % will be decomposed, and therefore, the emissions will be drastically reduced. In addition, the higher stable temperature is conducive to the improvement of parameter, quality and yield of steam, and realizing the waste recycling utilization. But the higher the temperature is, the shorter the residence time is, and it will increase the burden to the incinerator body and refractory materials. When the incinerator temperature is high enough, it should make the limitation for combustion speed. When the combustion temperature is too high, it will lead to high temperature slagging and corrosion, and increasing the incinerator body burden.

In summary, the temperature directly affects the combustion stability, burn-off rate, secondary pollutant emission, low temperature corrosion, high temperature slagging and production operation and so on, the incinerator temperature fluctuation will cause the change of burn-off rate, and it would lead slag combustion reduction rate exceed the standard and effects the steam production. The experience has shown that the incinerator operating at 850–950 °C range is rather less in flue gas pollutant emissions, and the slag combustion reduction rate is also minimal and stable.

4 Intelligent Control Strategy: Model and Algorithm

4.1 Intelligent Control Strategy

Control strategy selection should follow the principle that must be matched with cybernetics characteristic of controlled process. Aimed at the special cybernetics characteristics of combustion process such as uncertainty, nonlinearity and time-varying and so on, it is clear that the conventional PID control strategy is not desirable. Whether the classical control or the modern control of multi-input multi-output theory, all of them are the control strictly based on the mathematical model, and limited to the analysis and control of the linear process, they almost do not deal with the uncertainty, nonlinearity and time-varying process. Intelligent control is a cross discipline of control theory, artificial intelligence and computer science, the intelligent control strategy can solve the control puzzle that The conventional control strategy is difficult to control, and it provides more option space for control strategy. Aimed at the specific characteristics of the controlled process, it can adopt the control strategy that matches the cybernetics characteristics of controlled process, such as fuzzy logic control, expert control system, artificial neural network control and human simulated intelligent control (HSIC) [5] strategy and so on, and achieved a better control effect. In particular, it is noteworthy that the basic idea of human simulated intelligent control is the imitation of human and artificial intelligence, simulates the control structure and control behavior of human beings from macroscopical hierarchy, and therefore, the studied objective is the control mechanism of human beings itself. With the help of machines (controller), it realized the imitation of human control behavior, and had undergone from the qualitative reasoning process of "cognition to judgment" to quantitative control process of "judgment to operation". Therefore, the aim of the study is not the object (process), but it is the controller itself that is how to simulate the control structure and behavior. In essence, the intelligent control system is a sort of isomorphic transformation in body control system, and its basis is low layer control. This strategy has the following characteristics, feature recognition and feature memory on line, multimode control of combined open loop with closed loop and combined quantitative control with qualitative decision, application of direct viewing-type and heuristic reasoning logic. By means of production rule-based artificial intelligence, it can describe the intuitionistic reasoning logic. It can make full use of control theory, based on the classical control theory it can realize the control combined open loop with closed loop, based on combination of qualitative decision and quantitative control it can actualize the multimode control. Compared with other intelligent control strategy, the human simulated intelligent control strategy has outstanding advantages, and therefore it has been successfully applied more and more widely. The above is the reason using the control strategy for municipal solid waste incineration process.

4.2 Control Model

Any of a control system is composed by two parts of controlled process and controller, and the generalized control model is shown as in Fig. 1. In which, the process deviation $e(t) = r(t) - y(t)$, $r(t)$ is the process input, $u(t)$ is the controller output, $y(t)$ is the process output. The process deviation e, deviation change rate $\dot{e}$ and time t constitute a information space of solving the control problem. Obviously, whether it is constant value control and servo control, the control target is to make the fact that when time t goes to infinite the following must hold, and namely $e(t) = \dot{e}(t) = 0$.

In the information space, the trajectory of process state movement not only reflects all dynamic information of the process, but also reflects the effect of the control on the process. In the human being controller, the experienced operator is not based on mathematical model control, according to the inference of accumulated experience and knowledge (such as appearing the feature information in the process of dynamic response) sets or changes the control strategy online. In the information space, all the information feature sets reflect the dynamic characteristics of the process, and constitute the generalized mathematical model (feature model) of human simulated intelligent control. The feature model is the basis of feature identification, and it is the priori knowledge that the controller should own. For example, when $e \cdot \dot{e} > 0$ or $e = 0$ and $\dot{e} \neq 0$, then the absolute value e of process deviation demonstrates an increasing trend, and when $e \cdot \dot{e} < 0$ or $\dot{e} = 0$, then the absolute value e of process deviation demonstrates an decreasing trend, and With the change of time,the process deviation would automatically tend to return to zero. Thus, when $e = 0$ and $\dot{e} \neq 0$ or $e \cdot \dot{e} > 0$, selecting the proportion control mode is desirable, when $\dot{e} = 0$ or $e \cdot \dot{e} < 0$, selecting hold control mode is desirable, and both of them can make the process deviation automatically tend to return to zero.

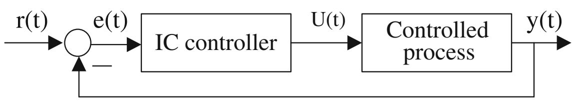


Fig. 1 Generalized control process model

$$u = \begin{cases} K_p e + kK_p \sum_{i=1}^{n-1} e_{m,i} & (e \cdot \dot{e} > 0 \cup e = 0 \cap \dot{e} \neq 0) \\ kK_p \sum_{i=1}^{n} e_{m,i} & (e \cdot \dot{e} > 0 \cup \dot{e} = 0) \end{cases} \tag{4}$$

In which, K_P is a proportional coefficient, $_k$ is an Inhibition coefficient, $e_{m,\,j}$ is the *j*th peak value of deviation, and the meaning of other symbol is the same as the previous. The expression (4) is an operation control algorithm. In the algorithm, based on the feature recognition relation between $\dot{e}$ and e, it can judge the dynamic process characteristics to be which motion state in two different modes, and then it can determine the control mode. Here, it not only owns the qualitative decision process (identification of motion state and selection of corresponding control mode), but also it owns the quantitative control (output of a specific control mode).

Combined with incineration process analysis, the temperature control is maintained in a given range from 850 to 950 °C, and the given value is set to 900 °C. When the temperature change rate significantly becomes higher, it explains that the waste calorific value is larger or unburned waste amount is larger, and vice versa. In order to ensure the stability of waste incineration, here it adopts two-dimensional intelligent controller, namely the double input and double output controller. The input variables are the temperature deviation and its change rate, and the output control quantities are the feeding time and stop feeding time. It can adopt the production rules of 'If… Then…'describes the control algorithm, and aimed at the specific process, the control algorithm can be summarized as shown in Table 1.

In Table 1, e is process deviation, K_P is proportional gain, K_1 is the gain amplification coefficient and $K_1 > 1$, K_2 is the inhibition coefficient and $0 < K_2 < 1$, $P_{O(n)}$ is the *n*th needed keeping value of control output P, and $P_{O(n-1)}$ is the $(n-1)$th needed keeping value of control output P. M_1, M_2 is respectively the deviation threshold and $M_1 > M_2$, and $e_{m,n}$ is the *n*th extremum of deviation.

Table 1 The table of HSIC algorithm

No.	If condition hold				Tlien output is	Mode
1	$\lvert e_n \rvert > M_1$				FFH or 00H	Switch
2	If	$e_n \cdot \Delta e_n > 0$ or $\Delta e_n = 0,\ e_n \neq 0$	and	$\lvert e_n \rvert \geq M_2$	$P_{0(n-1)} + K_1 K_{Pe_n}$	Proportion
				$\lvert e_n \rvert < M_2$	$P_{0(n-1)} + K_{Pe_n}$	
3	$e_n \cdot \Delta e_n < 0,\ \Delta e_n \cdot \Delta e_{n-1} > 0,\ e_n = 0$				$P_{0(n)} = P_{0(n-1)}$	Hold$_1$
4	If	$e_n \cdot \Delta e_n < 0$ $\Delta e_n \cdot \Delta e_{n-1} < 0$	and	$\lvert e_n \rvert \geq M_2$	$P_{0(n-1)} + K_1 K_2 K_{Pe_{m-n}}$	Hold$_2$
				$\lvert e_n \rvert < M_2$	$P_{0(n-1)} + K_2 K_{Pe_{m-n}}$	

5 Simulation Experiment and Its Analysis

In order to validate the effectiveness of control strategy and algorithm, under the model of BMCR boiler load conditions it makes the simulation experiment. It adopts the step disturbance to obtain the dynamic response characteristics of process output y(t). Taken PID control for reference, it studied robustness of HSIC control strategy and algorithm. After determining the physical structure model of data in the experiment site, the simulation of the waste incineration load is made using Table 1 algorithm, and the Fig. 2 is the simulation test results of the process in step input amplitude at 50. In order to ensure the stability of the initial steady state value in simulation, and the pulse disturbance is joined at 15th second. In Fig. 2, the disturbance pulse width of both Fig. 2a, b is 5 s.

The above content of simulation experiment includes that ① the main steam pressure and air volume keeps the constant, and the fuel volume makes the increasing disturbance. ② the main steam pressure and air volume keeps the constant, and the fuel volume makes the decreasing disturbance. ③ the fuel volume keeps the constant, the air volume makes the increasing disturbance. ④ the fuel volume keeps the constant, the air volume makes the decreasing disturbance. ⑤ the main steam pressure keeps the constant, and the fuel moisture makes the increasing disturbance. ⑥ the main steam pressure keeps the constant, and the fuel moisture makes the decreasing disturbance.

From comparative analysis of the simulation results, it can be seen that the HSIC has good control effect and anti jamming ability. In the case of large disturbances, it can quickly eliminate interference, the output deviation is controlled in the minimum range, and the output control of y(t) can quickly tend to smooth, therefore, it is propitious to improve the service life of actuating mechanism and the process stability. Compared with the conventional PID control, the HSIC shows a lot of advantages such as the fast response speed, short transition time, stable and smooth response curve, steady-state minimum deviation, strong anti-interference ability, good control effect and strong robustness and so on.

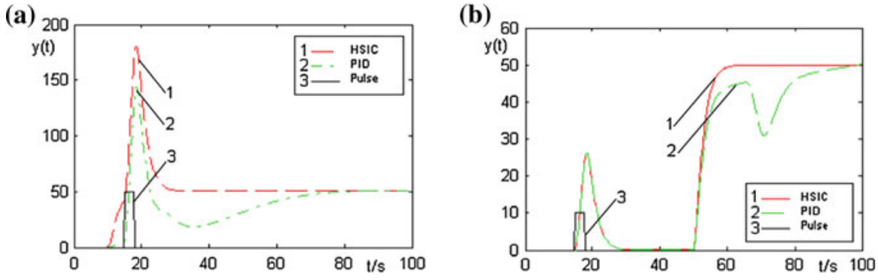


Fig. 2 Step response curve under different disturbance amplitude. **a** Response in amplitude at 50. **b** Response in amplitude at 10

6 Conclusions

Aimed at the existent problem in operation process of municipal waste power plant, the above explored the intelligent control strategy of incinerator combustion process. In the actual operation, it does not always rely on theoretical model to make the control, but with the help of the experience summed up the human being control system, it makes the human simulated intelligent control. The simulation results of pulse disturbance have verified the strong robustness of human simulated intelligent control. Although it only makes a qualitative analysis, the superiority of the HSIC strategy has shown that it is the available and effective to be applied in municipal waste incineration process.

References

1. Pei, Y.L., Yang, Z., Yang, X.Y.: An intelligent-fusion based control strategy of city garbage incineration. J. Liaoning Tech. Univ. (Nat. Sci.) **29**(5), 803–806 (2010)
2. Hua, X.G., Yang, Z.: The strategy of intelligent control for municipal waste mcineration system based on DCS. Program. Controll. Fact. Autom. **2006**(9), 123–128 (2006)
3. Ding, L.W., Li, F.R.: Technology features and pollution control strategy of MSW incineration power generation. Zhejiang Electr. Power **2010**(2), 55–58 (2010)
4. Xu, H.: Control strategy of municipal refuse incineration power. Energy Eng. **2005**(5), 38–41 (2005)
5. Li, Z.S., Tu, Y.Q.: Human Simulated Intelligent Controller. National Defense Industry Press, Beijing (2003)

Intelligent Control of Forming Process in Complex Work-Pieces of Plastic Cement

Hong-jian Tao and Hu Xu

Abstract Plastic materials own complex chemical composition, aimed at the puzzle of being difficult to actualize precisely control in temperature holding time, injection pressure and process temperature of the plastics work-piece forming process, the paper explored the strong robustness strategy based on multi-modal control. This paper made the anatomy in control puzzle of existent process, summarized up the process characteristics in cybernetics, based on the comparative research for control strategy, explored the control strategy of strong robustness based on human simulated intelligence, and constructed the multi-modal control algorithm of the plastics work-piece forming process. Taking the control of process temperature as an example, the experimental simulation validated the good dynamic and static control quality of strong robustness control strategy proposed in this paper. The experimental result shows that it is reasonable and feasible to strong robustness control strategy explored in the paper.

Keywords Pieces of plastic processing · Forming process · Robust control strategy · Multi-modal control

1 Introduction

The plastic work-piece is widely used in the engineering field, but the physical and chemical properties of product is affected by many factors constrained such as compound formulation, structure size, molding technology and processing environment and so on. In the production process, currently it mainly depends on the

H. Tao (✉)
School of Information Engineering, Chongqing Industry Polytechnic College,
Chongqing 401120, China
e-mail: 108405999@qq.com

H. Xu
College of Automation, Chongqing University, Chongqing 400044, China
e-mail: 1597725908@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_31

operator's experience and skill to conduct the control of production process because of the complexity production technology, and therefore, the stability of product quality is subject to the subjective initiative of operator to a great degree. In view of the production process control being difficult to guarantee the stability of product quality, and in order to avoid being affected by artificial factors, the key to improve the stability of the product quality is to adopt the correct control strategy [1, 2]. So, this paper makes some discussions on the control strategy of plastic work-piece forming process.

2 Control Puzzle of Forming Process

Under the specified conditions, the forming control is a physicochemical process that the plastic material and various additive arises the chemical cross-linking reaction. Except the control process is affected by the factors such as product formula, material properties, product structure, technology conditions and so on, and the key factors, those have a decisive influence on the product properties, are forming temperature, temperature-retaining time and molding injection pressure in the forming control process. For example, a certain pressure can promote the plastic material flow in the pressed film and remove the air bubbles, and ensures the density of products. The pressure depends on the product structure, the nature of compound and technology process conditions, and if the mobility of plastic material is poor, then the pressure should be greater for thick products on the structure and complex shapes. So, it can make the plastic material be able quickly to fill the whole cavity, removes the air bubbles, and improves the adhesion and flexibility etc. between the product layer and the metal layer or among the cloth layers. Because the plastic material has no chemical molecule of fixed type, it belongs to the high polymer material. Even if it is the same kind of raw materials, the molecular weight also has significant difference, it may be different a hundred times or even thousands of times, and full of uncertainty. In the forming process, the heat transfer is achieved by thermal radiation and convection. The distribution of temperature field is not uniform, and it describes the heat transfer relationship only by differential equations or partial differential equation. There is obviously the heat transfer lag, it belongs to the non-steady state system, and there is a strong nonlinear obviously. As the heat transfer of plastic material and heat capacity is changeable, its heat property is changed with chemical reaction, and so the heat transfer lag is unknown and time-variant. The plastic material is a poor conductor of heat, the heat conduction is related to the temperature etc., and therefore, the heat parameters are randomness, unknown, diversity and time-variant. In the forming technology of molded plastic products, the key technology affecting the quality stability of forming product is to actualize the optimal control of the forming time, injection pressure and process temperature, and this is currently the control puzzle of forming process control faced [3]. As to what control strategy should be taken, it deals with the research on cybernetics characteristic of product forming process.

3 Process Characteristic and Control Strategy

3.1 Process Characteristic

Based on the analysis of control puzzle mentioned above forming process of work-piece, it can be summarized up the cybernetics characteristics of forming process [4]. ① Uncertainty of process parameter. Because the plastic materials have no chemical molecule formula fixed type, it belongs to the high polymer material. The forming technology has no obvious regularity, and the process is closely related to the factors such as plastic material formulation, technology conditions, product structure and properties etc. The uncertainty results in randomness, distribution, time-varying and unknown, and it is to say that the process parameters are just an estimate of the value. ② Large inertia and pure time-lag of process. The plastic cement is not a good conductor of heat transfer, the temperature field distribution is not uniform, and there is obviously large inertia and large pure time lag in the heat transfer of controlled process. ③ Process presents strong nonlinear. The heat conduction equation itself of describing the process characteristics is nonlinear differential equation or partial differential equation, and therefore, the whole forming process shows a strong nonlinearity. ④ Time-varying and time lag unknown of process. In the process of chemical reaction, the properties of the material are continuously changing, and the factors of many uncertainties result in time-varying and time lag unknown of the controlled process with time lag in plastic material forming. ⑤ The environmental disturbance of forming process has unpredictability and diversity. Facing the process characteristics in cybernetics, it is not possible to obtain the expected control effect by using the conventional control strategy, and therefore, it is necessary to find a new control strategy.

3.2 Control Strategy

It is not desirable to the normalized control based on classical and modern control theory because of its mathematical modeling difficulty being too big, and the engineering practice has proved that it is impossible to achieve optimal control by means of the conventional control strategy. The control strategy can be generally adopted as intelligence control strategy. Such as fuzzy control strategy, it can use the fuzzy set theory to make the fuzzification by linguistic variables, and after the fuzzy reasoning, the process can be controlled according to the fuzzy control rules. Although there are many successful application cases, and in the face of uncertainty in process control, it is incapable of action, and also the control effect is not ideal. The expert control system based on knowledge can be able to handle a variety of accurate, fuzzy, qualitative and quantitative information, but for the uncertainty of the process, it is not easy to the acquisition of relevant transcendental knowledge, and therefore, it is very difficult to provide the control pattern to be suitable for process control. What it is worth to pay the particular concern is the human simulated intelligent control strategy. It is to study on control mechanism and control structure of the human being controller itself, and simulates the control behavior of human. Essentially it is the "homomorphism" transformation of human control system. Adopting generalized characteristics model based on knowledge to actualize the control, it not only considers dynamical characteristic of process, but also researches on static characteristic of process, and it studies on control strategy, control pattern and control algorithm from system hierarchy, and therefore, the control represents the characteristics attracted people's attention that it is high in control precision, short in transient process time, fast in response time, none in overshoot, strong in robustness, easy to fuse into the actual experience, skill and wisdom of operator etc. HSIC (Human Simulated Intelligent Controller) owns the following advantages. The strategy can make the online characteristic recognition and has the characteristic memory function, adopt the principle of combining both qualitative and quantitative control using the integration of open and closed loop control, simulate the control behavior of control expert by means of heuristic and intuitive reasoning method, and actualize the multi-modal control based on human simulated intelligence. Because the control strategy takes fully the account of more things in requirements of dynamic and static quality index, it integrates various performance demands such as response speed, smoothness, robustness and accuracy, and as a whole it is closer to the actual project application. The successful engineering practice has proved the superiority of the control strategy. Therefore the paper selects the control strategy based on human simulated intelligence.

4 Control Model and Control Algorithm

4.1 Control Model

Generalized control model based on human simulating intelligence is an intelligent system in essence [5]. First of all, it makes the integration to user index and the of knowledge base and fuses a generalized knowledge set, then it takes the knowledge set, reasoning mechanism and control rule to be fused into a human simulated intelligent controller. The structure of generalized control model is shown as in Fig. 1, and in which, r(t), e(t), u(t), y(t) is respectively represents the process input, process error, output of controller and output of controlled process.

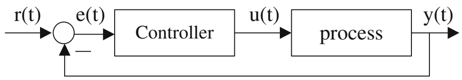


Fig. 1 Generalized control process model

4.2 Control Algorithm

Taken the original algorithm of HSIC as an example, its control algorithm is as the following.

$$u = \begin{cases} K_p e + kK_p \sum_{i=1}^{n-1} e_{m,i} & (e \cdot \dot{e} > 0 \cup e = 0 \cap \dot{e} \neq 0) \\ kK_p \sum_{i=1}^{n} e_{m,i} & (e \cdot \dot{e} > 0 \cup \dot{e} = 0) \end{cases}$$

In which, u is the output of controller, K_p is a proportional coefficient, k is an inhibition coefficient, e is the system error, $\dot{e}$ is the error change rate, $e_{m,j}$ is the ith peak value of the error. Based on the above original control algorithm, and aimed at the characteristic of the special forming process in complex work-pieces of plastic cement, it can further summarize the human control experience, and establish a set of control rules matched with the cybernetics characteristic of forming process. The specific rules are as the following.

If $e > \beta R$, Then $u_n = u_m$.
If $e < -\beta R$, Then $u_n = -u_m$.
If $|e| < \delta_1$ and $|\dot{e}| < \delta_2$, Then $u_n = u_{n-1}$.
IF $|e| > m \cdot R$ and $e \cdot \dot{e} > 0$, Then $u_n = u_p$.
If $e \cdot \dot{e} < 0$ and $|e/\dot{e}| > a$, Then $u_n = \alpha p_1$.
If $e \cdot \dot{e} < 0$ and $|e/\dot{e}| > b$, Then $u_n = p_1 + K_d \dot{e}$.
If $e \cdot \dot{e} < 0$ and $b \leq |e/\dot{e}| \leq a$, Then $u_n = p_1$.
If $e \cdot \dot{e} \geq 0$ and $|e| \in (\delta_1, \theta_1), |\dot{e}| \in (\delta_2, \theta_2)$,
Then $u_n = P_1 + K_{p2} e + K_{i2} \sum e_j$.
If $e \cdot \dot{e} < 0$ and $|e| \in (\delta_1, \theta_1), |\dot{e}| \in (\delta_2, \theta_2)$,
Then $u_n = P_1 + K_{p2} e$.
If $e \cdot \dot{e} \geq 0$ and $|e| > \theta_2$,
Then $u_n = P_1 + K_{p1} e + K_{i1} \sum e_j - K_d \dot{y}$.

In the above rules, u_n is the nth output of controller, u_m, is a output keeping value related to the input change amount ΔR, up is a Forced retention value. e, $\dot{e}$ is respectively the system error and its change rate, $\dot{y}$ is the change rate of system output. $e_{m,j}$ is the jth extremum of the error, r is a weighted factor of extremum value, and it can be online correction. K_{p1}, K_{p2}, K_{i1}, K_{i2}, K_d is respectively the proportional, integral and differential gain. β is a switching factor. α, a, b is respectively different constant, and it is determined by empirical rule in knowledge set. R is the setting value. $\delta_1, \delta_2, \theta_1, \theta_2$ is respectively the allowed error and error rate range. $\sum_j e_j$ is the cumulative error during period of $e \cdot \dot{e} \geq 0$. $p_1 = r \sum_{i=1}^{l} e_{m,i}$ is a recent retention value of controller.

5 Process Simulation and Experimental Verification

5.1 Process Simulation

In order to verify the control performance, it can inspect the performance in the aspect of process parameter change in process simulation, and taking the conventional PID control strategy for reference, it makes the comparison of process response by above specific algorithm in this paper, and observes the change of performance index in rise time, adjusting time, whether overshoot and steady-state control accuracy etc.

A two-order process model with time lag is as the following.

$$G(s) = \frac{4.134 \times e^{-10s}}{(s+1)(2s+1)}$$

The change of cybernetics characteristic in plastics work-piece forming process can be equivalent to the effect of process parameter change, and make the process simulation. Transforming the original model G(s) = $4.134e^{-10s}/(s + 1)(2 s + 1)$ as the G'(s) = $12e^{-2s}/(1.2 s + 1)(5 s + 1)$, and obviously, all the parameters of the process model are changed. Under this situation, the unit step response of process G (s) and process G'(s) is respectively shown as in Fig. 2 and Fig. 3. From the analysis of Figs. 2 and 3, it can be seen that in the aspects of performance index such as process overshoot, control precision, response smoothness, adjusting time and rise time etc., the control strategy of HSIC represents the better control quality and strong robustness, and the control of HSIC obviously excels PID control strategy.

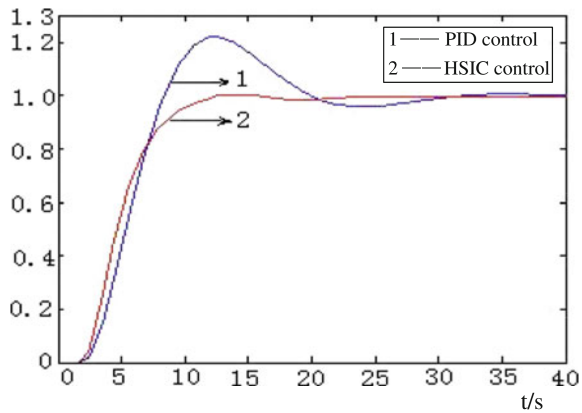


Fig. 2 Unit step response of original model

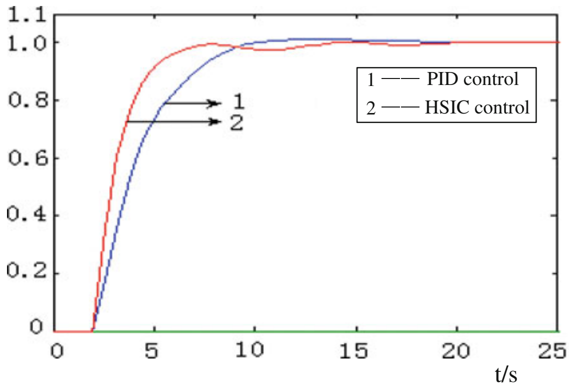


Fig. 3 Unit step response of parameter change

5.2 *Experimental Verification*

Taking the forming control of plastic bearing work-piece as an example, the forming control process is adopted respectively by using the PID and HSIC control, before and after the system reformed, the desired pressure is set to 16 MPa. The change range of original system pressure is from 16.50 to 15.12 MPa, and the change range of existent system pressure is from 16.14 to 15.84 MPa. Taking 100 mS as the sampling period, it makes the sampling and recording for two pressure system, the sampling time takes 10 min, and the mean value, variance and range are shown as in Table 1. Data in the Table 1 show that the HSIC control is obviously better than PID control.

Before and after the system reformed, the desired temperature is set to 165 °C. The change range of original system temperature is from 152 to 173 °C, and the change range of existent system temperature is from 163 to 166.5 °C. Taking 100 mS as the sampling period, it makes the sampling and recording for two pressure system, the sampling time takes 10 min, and the mean value, variance and range are shown as in Table 2. Data in the Table 2 show that the HSIC control is obviously better than PID control.

Table 1 Contrast between pressures

System	Stat.		
	Average	Variance	Range
Original	15.82	0.62	1.38
Existent	15.97	0.12	0.30

Table 2 Contrast between temperatures

System	Stat.		
	Average	Variance	Range
Original	163.2	8.0	21.0
Existent	164.7	1.1	3.5

5.3 Result Analysis

The simulation in Figs. 2 and 3 shows that when the process parameters change, the control algorithm of HSIC still has good control quality, and compared with PID control algorithm, it has very obvious advantage. Since the algorithm can make the adjustment in real time according to the error signal of process, it improves the control action for process, and therefore, the control HSIC of strong robustness has the absolute advantage for the forming process control of complex plastic work-pieces. In addition, the engineering experiment data in Tables 1 and 2 shows that the control algorithm of HSIC is more ideal, and it has stronger engineering applicability in the aspects of control quality such as control precision, average value, variance and range.

6 Conclusions

The key of forming process control is that the parameters of process temperature, injection pressure and keeping time should be better to be controlled. In view of the specialty in cybernetics characteristics, and due to the effects of many uncertainty factors, it is very difficult to control. The proposed control strategy based on HSIC has very strong robustness, the simulation experiment and engineering practice show that it is feasible and effective to the control of plastic forming process.

References

1. Xiong, X.S.: Strong robust control strategy of complex combustion process with uncertainty. J. Chongqing Univ. Technol. (Nat. Sci.) **27**(6), 106–110 (2013)
2. Wang, J., Xiao, C.: Multi-modal simulated intelligent control strategy for complex process. Microelectronics **43**(2), 221–224 (2013)
3. Pei, Y.L.: Intelligence based fusion control strategy for complex vulcanizing process. Hydromechatronics Eng. **41**(6), 92–96 (2013)
4. Wang, J.: Intelligence based fusion aeration control strategy of polluted water treatment. J. Southwest Univ. (Nat. Sci. Ed.) **34**(7), 120–124 (2012)
5. Li, Z.S., Tu, Y.Q.: Human Simulated Intelligent Controller. National Defense Industry Press, Beijing (2003)

A Study on Risk Factors Impact on the Efficiency of Commercial Banks in China—Based on Fuzzy DEA Model

Xin Liu Bao-zhu Shi and Xiang Zou

Abstract Based on the study of risk factors impact on the efficiency of commercial Banks in China, is mainly to make use of the non-performing loan ratio, capital adequacy and overdue loans to risk adjustment of input and output, and then uses the fuzzy SBM and Super-SBM model to estimate the efficiency. Results show that: the influence of risk factors on the 13 commercial banks in 2011 has varying degrees. The less risk faced by banks, the smaller influence on the efficiency affected by the risk factors. Compared with the state-owned commercial banks, the risk factors have a higher influence on the joint-stock commercial banks.

Keywords Commercial banks · Efficiency · Fuzzy DEA · Risk adjustment

1 Introduction

Banking industries in China have changed intensively after reform and opening-up. Many commercial banks transformed into joint-stock enterprises and appear on the market. The factors of external sources of finance and economic crisis all have influence on the development of banks. With the rapid development of the banking industries, the problem of poor efficiency is serious and it also brings about many questions—such as poor operation management level, lack of innovation about financial products, less competition, high bad loan ratio, nonstandard capital risk management and so on. All of these will influence the development of bank industries.

Indian commercial banks have been classified into two ownership groups (publicly owned and privately owned) by Sathye [1]. Then the measurement of efficiency is done using data envelopment analysis. The study shows that the efficiency of private sector commercial banks is lower than that of public sector banks. Hsiao et al.

X.L.B. Shi (✉) · X. Zou
School of Economics, Dongbei University of Finance and Economics,
Liaoning 116025, China
e-mail: liuxin060@sina.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_32

[2] used a fuzzy super-efficiency slack-based measure DEA to analyze the efficiency of 24 Taiwan's commercial banks with vague characteristics. The study shows that the fuzzy super-efficiency slack-based measure of efficiency can not only characterize uncertain input variables and output variables, but also have a higher capability to evaluate bank efficiency than the conventional Fuzzy DEA approach. Eslami et al. [3] put fuzzy probabilistic constraints into classic DEA, and analyzed the data sources consist of monthly reports of 20 Iranian bank branches over a period of May 2008–February 2009. The study shows that the fuzzy super-efficiency slack-based measure of efficiency can not only characterize uncertain input variables and output variables, but also have a higher capability to evaluate bank efficiency than the conventional Fuzzy DEA approach. Compared with bank efficiency studies abroad, domestic studies started off not so early, which, moreover, lag behind in research methods together with philosophy to some degree. For example, Huang Xian et al. [8] gave an study on the technical efficiency of 13 commercial banks in China during the period 1998~2005 by applying the three-stage DEA model. When Tang Zhuangzhi [9] measured the efficiency of Chinese banks by applying DEA model, he used Factor Analysis Method to screen out suitable input-output variable, moreover put risk factor into it. Hu Han and Song Yuanliang [10] used SFA model to study Chinese commercial bank's efficiency and risk factor's influence on it. From the efficiency value, we will find that Chinese bank's total efficiency level is low during sample period. From that, state-owned commercial banks' efficiency is higher than joint-stock commercial banks' efficiency. The study of risk factors shows that loan-deposit ratio and capital adequacy ratio have positive correlation relationship with bank's efficiency, in addition, bad loan ratio has negative correlation relationship with bank's efficiency.

Above-mentioned available literature about commercial bank's efficiency all hypothesized that risk is neuter or that risk factors have so little influence on bank's efficiency that we can ignore it. So the measurement of bank's efficiency deviated from the true value. However, the risk itself is an uncertain message, and it also brings about capital's fluctuation and income's instability. It is suitable to show this uncertainty by applying vague number which is flexible structure. So we try to study risk factors influence on Chinese commercial banks' efficiency by using fuzzy DEA model.

2 Fuzzy DEA Model

From the econometrics view, conventional DEA method is a pure linear programming process, moreover, compared with classical DEA, fuzzy DEA obviously increase calculate amount because fuzzy DEA fuzzify the index which is quantification by using L-R vague number and triangle vague number. In addition, the efficiency from conventional DEA is simple maximum value, but the efficiency from fuzzy DEA is a minimum one among the maximum value, and it is a precise maximum value.

2.1 Basic Definition of Fuzzy Number

Definition 1 Assuming X is an Universe of Discourse, $\tilde{A}$ a fuzzy set in X, is shown by a real-valued function which is subject to $X, \mu_{\tilde{A}}: X \to [0, 1], x \to \mu_{\tilde{A}}(x)$, for $x \in X$, the function $\mu_{\tilde{A}}$ is a subjection function of $\tilde{A}$. The function value of $\mu_{\tilde{A}}(x)$ is called the membership of x for $\tilde{A}$, denoted as: $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | x \in X\}$.

Definition 2 Assuming R is real numbers set, and $\tilde{A}$ is fuzzy set, if for any real number $< y < z$, there are $\mu_{\tilde{A}} \geq min\left[\mu_{\tilde{A}(x)}, \mu_{\tilde{A}(z)}\right]$, so we call $\tilde{A}$ as convex fuzzy set on R.

Definition 3 Assuming the domain of discourse of $X = R$ is real number set, if

(1) $\tilde{A}$ is a convex fuzzy set on X;
(2) $\exists\, a_1, a_2 \in X$ and $a_1 \leq a_2$, subject to $\forall x \in [a_1, a_2], \tilde{A}(x) = 1$, $\lim_{x \to +\infty} \tilde{A}(x) = 0$, then $\tilde{A}$ is a fuzzy number on X.

Definition 4 If $\tilde{A} = [a_1, a_M, a_2]$, $0 \leq a_1 \leq a_M \leq a_2$, then call $\tilde{A} = [a_1, a_M, a_2]$ the triangular fuzzy number. The membership function of $\tilde{A}$ can be represented (Van Laarhoven and Pedrycz [11]):

$$\alpha = \mu_{\tilde{A}}(x) = \begin{cases} \frac{x - a_1}{a_M - a_1}, a_1 \leq x \leq a_M \\ \frac{x - a_2}{a_M - a_2}, a_M \leq x \leq a_2 \\ 0, \quad others \end{cases} \tag{1}$$

This article mainly use input-output variable to construct triangular fuzzy number, transforming triangular fuzzy number to interval number by applying α–cuts, then getting the result by using interval DEA model.

2.2 Constructing Fuzzy Number

By applying classical BCC model, we not only can measure the efficiency of every decision unite, but also can get the input differential value of every DMU. We use SFA model to separate the influence of risk factor and stochastic interference factor on input differential value, then have a risk-adjustment on input variable, at last, do a fuzzy DEA analysis on input variables’ fuzzy number.

If there are n decision unites, every decision unite have m input methods, and there are p risk variables having influence on input differential value. The input differential regression function is:

$$S_{ij} = f^i\left(z_j;\beta^i\right) + v_{ij} + u_{ij}, i = 1, 2, \ldots m; j = 1, 2, \ldots, n \quad (2)$$

In function (2), S_{ij} represent i th input's difference of j th decision making unit; $z_j = (z_{1j}, z_{2j}, \ldots, z_{pj})$ represent p exogenous risk variable, β^j is unestimated parameter of risk explanatory variable; $f^i(z_j;\beta^i)$ shows the function relationship between risk variable and input difference. v_{ij} is ith input's random term of jth DMU, u_{ij} is risk inefficiency's nonnegative random variable of ith input of jth DMU. We assume v_{ij} complying with normal distribution $(0, \sigma^2_{vi})$, while u_{ij} obeying truncated normal distribution $N^+\left(\mu^j, \sigma^2_{ui}\right)$, in addition v_{ij} and u_{ij} are mutual independence.

By using Frontier Version software, we estimate the unknown parameter of function (2) by applying Maximum likelihood estimate. Meanwhile, we can estimate $\hat{E}[u_{ij}/v_{ij} + u_{ij}]$ by using the method of Jondrow et al. [4]. Based on this, the value estimator of v_{ij} is:

$$\hat{E}\left[\frac{v_{ij}}{v_{ij} + u_{ij}}\right] = S_{ij} - z_j\beta^i - \hat{E}[u_{ij}/(v_{ij} + u_{ij})] \quad (3)$$

This paper uses estimated parameter to have a risk-adjustment on every DMU, the principle is to adjust all of the DMU into a same risk condition, at the same time, consider of the influence of stochastic disturbance term, so we can get the efficiency which reflect decision unite in risk condition. The adjustment expression is:

$$x^U_{ij} = x_{ij} + \left\{max_j\left[z_j\beta^i\right] - z_j\beta^i\right\} + \left\{max_j[\hat{v}_{ij}] - \hat{v}_{ij}\right\} \quad (4)$$

$$x^L_{ij} = x_{ij} - \left\{z_j\beta^i - min_j[z_j\beta^i]\right\} - \{\hat{v}_{ij} - min_j[\hat{v}_{ij}]\} \quad (5)$$

Among it, x_{ij} is section i's real value of section j of DMU, β^i is the estimator of risk explanatory variable, $\hat{v}_{ij}$ is the estimator of stochastic disturbance term.

In Eq. (4), the first bracket represent an adjustment of risk variable, and it makes all sample banks face the same great risk condition. Among Eq. (5), the first bracket also represent an adjustment of risk variable, but it makes all sample banks face the same low risk. The second bracket of the two equations depicts the adjustment of stochastic disturbance term, and all the sample banks are confronted with a same random error because of it. x^U_{ij} represents input which has been adjust by enlarging risk, x^L_{ij} represents input which has been adjust by decreasing risk, then we can build up triangular fuzzy number as $\left[x^L_{ij}, x_{ij}, x^U_{ij}\right]$ for every input variable.

2.3 SBM Model and It's Solution in Fuzzy Condition

Fuzzy DEA model expands based on classical DEA models, so different classical DEA model has different fuzzy DEA model. This article chooses non-radial SBM model to study, so conventional CCR model and BCC model have incomplete measurement. Their model either is input-oriented or output-oriented, so we can only get input efficiency or output efficiency of DMU_s while can not get the measurement of input and output efficiency at the same time, and their efficiency index miss the slack quantity of nonzero input-output, therefore, they can't get the inefficiency of all DMU_s. While non-radial SBM model directly put the slack quantity input and output into objective function and constraint function, so on the one hand, it solves the measurement problem of slack quantity of input and output, on the other hand, it gives the efficiency based on input and output, so the estimate of efficiency about fuzzy DMU_s is more perfect and more reasonable.

2.3.1 Non-radial SBM Model

SBM model, which is proposed by Tone [5], is used to measure non-radial efficiency of decision making unite. We assume there are n DMU_s, each DMU with m inputs, and s outputs. Based on this description, the efficiency evaluation model could be expressed:

$$\begin{aligned} \text{Min} \quad & \rho_k = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} \frac{s_i^-}{x_{ik}}}{1 + \frac{1}{s}\sum_{r=1}^{s} \frac{s_r^+}{y_{rk}}} \\ \text{Subject to} \quad & x_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, i = 1, 2, \ldots, m \\ & y_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, r = 1, 2, \ldots, s \\ & \sum_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geq 0, j = 1, 2, \ldots, n \\ & s_i^- \geq 0, s_r^+ \geq 0 \end{aligned} \tag{6}$$

where (x_{ik}, y_{rk}) represents the kth DMU, the ith input, and rth outputs. The λ_j is represented by the kth DMU weighting for evaluating efficiency. The s_i^- is represented by input slack variable, while s_r^+ is represented by output slack variable. Model (6) is difficult to calculate, so we can multiply non-negative scalar variable q to the denominator a numerator, and assume that the objective function denominator is equal to 1. As such, Model (6) forms a simple linear programming similar to Model (7):

$$
\begin{aligned}
&\text{Min} && \tilde{\rho}_k = q - \tfrac{1}{m}\textstyle\sum_{i=1}^{m} \tfrac{S_i^-}{x_{ik}} \\
& && q + \tfrac{1}{S}\textstyle\sum_{r=1}^{S} \tfrac{S_r^+}{y_{rk}} = 1 \\
&\text{Subject to} && qx_{ik} = \textstyle\sum_{j=1}^{n} x_{ij}\lambda_j^{'} + S_i^-, i = 1, 2, \ldots, m \\
& && qy_{rk} = \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j^{'} - S_r^+, r = 1, 2, \ldots, s \\
& && \textstyle\sum_{j=1}^{n} \lambda_j^{'} = q, q > 0 \\
& && \lambda_j^{'} = q\lambda_j \geq 0, j = 1, 2, \ldots, n \\
& && S_i^- \geq 0, S_i^- = qs_i^- \\
& && S_r^+ \geq 0, S_r^+ = qs_r^+
\end{aligned} \tag{7}
$$

When the optimum solution of Model (7) $\tilde{\rho}_k^* = 1, S_i^{-*} = S_r^{+*} = 0,$ the efficiency of decision making unite is 1, and there aren't input excess and output shortfall, DMU_k is SBM's efficiency.

2.3.2 Fuzzy Expand and Solution of SBM Model

Among a DMU_s set, in order to represent fuzzy input and output variables, we used $\mu_{\tilde{x}_{ij}}$, $\mu_{\tilde{y}_{ij}}$ as their membership functions, such that $\tilde{x}_{ij}, \tilde{y}_{rj}$ are the fuzzy number of input and output, respectively. Usually, we assume all variables as fuzzy number, while if some variables are crisp data, then we can use degenerated member functions to describe them, such that this range will only have one value. Fuzzy SBM model is as follows:

$$
\begin{aligned}
&\text{Min} && \tilde{\rho}_k = q - \tfrac{1}{m}\textstyle\sum_{i=1}^{m} \tfrac{S_i^-}{\tilde{x}_{ik}} \\
&\text{Subject to} && q + \tfrac{1}{S}\textstyle\sum_{r=1}^{S} \tfrac{S_r^+}{\tilde{y}_{rk}} = 1 \\
& && q\tilde{x}_{ik} = \textstyle\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j^{'} + S_i^-, i = 1, 2, \ldots, m \\
& && q\tilde{y}_{rk} = \textstyle\sum_{j=1}^{n} \tilde{y}_{rj}\lambda_j^{'} - S_r^+, r = 1, 2, \ldots, s \\
& && \textstyle\sum_{j=1}^{n} \lambda_j^{'} = q, q > 0 \\
& && \lambda_j^{'} = q\lambda_j \geq 0, j = 1, 2, \ldots, n \\
& && S_i^- \geq 0, S_i^- = qs_i^- \\
& && S_r^+ \geq 0, S_r^+ = qs_r^+
\end{aligned} \tag{8}
$$

We assume $(\tilde{x}_{ij}), S(\tilde{y}_{rj})$ are $\tilde{x}_{ij}, \tilde{y}_{rj}$'s support sets, the $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ y α–cut is therefore defined as:

$$(X_{ij})_\alpha = \left\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geq \alpha\right\}, \forall j, i \tag{9}$$

$$(Y_{rj})_\alpha = \left\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geq \alpha\right\}, \forall r, j \tag{10}$$

Thus, fuzzy inputs and outputs can be represented in intervals at different α–cut levels $\{(\tilde{x}_{ij})_\alpha | 0 < \alpha \le 1\}$ and $\left\{(\tilde{y}_{rj})_\alpha | 0 < \alpha \le 1\right\}$. Based on the above, we can express (Eqs. (9) and (10) as Model (11) and Model (12)):

$$\begin{aligned}(X_{ij})_\alpha &= \left\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \ge \alpha\right\} = \left[(X_{ij})^L_\alpha, (X_{ij})^U_\alpha\right] \\ &= \left[min_{x_{ij}}\left\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \ge \alpha\right\}, max_{x_{ij}}\left\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \ge \alpha\right\}\right]\end{aligned} \tag{11}$$

$$\begin{aligned}(Y_{rj})_\alpha &= \left\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha\right\} = \left[(Y_{rj})^L_\alpha, (Y_{rj})^U_\alpha\right] \\ &= \left[min_{y_{rj}}\left\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha\right\}, max_{y_{rj}}\left\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha\right\}\right]\end{aligned} \tag{12}$$

This article assume input as triangular fuzzy number $X_{ij} = \left(x^L_{ij}, x^M_{ij}, x^U_{ij}\right)$, output as $Y_{rj} = \left(y^L_{rj}, y^M_{rj}, y^U_{rj}\right)$, its corresponding α–cut sets based on given α–level are:

$$(X_{ij})_\alpha = \left[x^L_{ij} + \alpha\left(x^M_{ij} - x^L_{ij}\right), x^U_{ij} - \alpha\left(x^U_{ij} - x^M_{ij}\right)\right] \tag{13}$$

$$(Y_{rj})_\alpha = \left[y^L_{rj} + \alpha\left(y^M_{rj} - y^L_{rj}\right), y^U_{rj} - \alpha\left(y^U_{rj} - y^M_{rj}\right)\right] \tag{14}$$

Moreover, we can define the DMU_k membership functions using Zadeh's (1978) extension principle, as seen in Model (15):

$$\mu_{\tilde{E}_k}(z) = \sup_{x,y} min\left\{\mu_{\tilde{x}_{ij}}(x_{ij}), \mu_{\tilde{y}_{rj}}(y_{rj}), \forall j, r, i | z = E_k(x, y)\right\} \tag{15}$$

where $E_k(x, y)$ is represented by the kth DMU efficiency score at a set of (x, y) by conventional SBM model and the $\mu_{\tilde{E}_k}$ is the minimum of $\left(\mu_{\tilde{x}_{ij}}(x_{ij}), \mu_{\tilde{y}_{rj}}(y_{rj})\right)$. From Model (15), we need $\mu_{\tilde{x}_{ij}}(x_{ij}) \ge \alpha, \mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha$ and at least one $\mu_{\tilde{x}_{ij}}(x_{ij})$ or $\mu_{\tilde{y}_{rj}}(y_{rj})$ equal to α, such that $E_k(x, y)$ satisfies $\mu_{\tilde{E}_k}(z) = \alpha$. Based on given α level, all α–cut sets form a nested structure, when $0 < \alpha_2 < \alpha_1 < 1$, there are: $\left[(X_{ij})^L_{\alpha_1}, (X_{ij})^U_{\alpha_1}\right] \subseteq \left[(X_{ij})^L_{\alpha_2}, (X_{ij})^U_{\alpha_2}\right]$, $\left[(Y_{rj})^L_{\alpha_1}, (Y_{rj})^U_{\alpha_1}\right] \subseteq \left[(Y_{rj})^L_{\alpha_2}, (Y_{rj})^U_{\alpha_2}\right]$, so $\mu_{x_{ij}}(x_{ij}) \ge \alpha$ and $\mu_{\tilde{x}_{ij}}(x_{ij}) \ge \alpha$ have a same domain and the same with $\mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha$ and $\mu_{\tilde{y}_{rj}}(y_{rj}) \ge \alpha$.

Based on the above formulas, we can find that $\mu_{\tilde{E}_k}$ has a lower and upper bound by α–cut, as seen in Model (16) and Model (17):

$$(\rho_k)_\alpha^U = \max_{\substack{(X_{ij})_\alpha^L \le x_{ij} \le (X_{ij})_\alpha^U \\ (Y_{rj})_\alpha^L \le y_{rj} \le (Y_{rj})_\alpha^U \\ \forall j, i, r}} \begin{cases} \min \quad q - \frac{1}{m}\sum_{i=1}^{m} \frac{S_i^-}{x_{ik}} \\ \textit{subject to} \quad q + \frac{1}{S}\sum_{r=1}^{S} \frac{S_r^+}{y_{rk}} = 1 \\ qx_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j^{'} + S_i^-, i = 1, 2, \ldots, m \\ qy_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j^{'} - S_r^+, r = 1, 2, \ldots, s \\ \sum_{j=1}^{n} \lambda_j^{'} = q, q > 0 \\ \lambda_j^{'} = q\lambda_j \ge 0, j = 1, 2, \ldots, n \\ S_i^- \ge 0, S_r^+ \ge 0 \end{cases} \tag{16}$$

$$(\rho_k)_\alpha^L = \min_{\substack{(X_{ij})_\alpha^L \le x_{ij} \le (X_{ij})_\alpha^U \\ (Y_{rj})_\alpha^L \le y_{rj} \le (Y_{rj})_\alpha^U \\ \forall j, i, r}} \begin{cases} \min \quad q - \frac{1}{m}\sum_{i=1}^{m} \frac{S_i^-}{x_{ik}} \\ \textit{subject to} \quad q + \frac{1}{S}\sum_{r=1}^{S} \frac{S_r^+}{y_{rk}} = 1 \\ qx_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j^{'} + S_i^-, i = 1, 2, \ldots, m \\ qy_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j^{'} - S_r^+, r = 1, 2, \ldots, s \\ \sum_{j=1}^{n} \lambda_j^{'} = q, q > 0 \\ \lambda_j^{'} = q\lambda_j \ge 0, j = 1, 2, \ldots, n \\ S_i^- \ge 0, S_r^+ \ge 0 \end{cases} \tag{17}$$

Models (16) and (17) are two-stage mathematical programming model, so we should apply the formula of Model (16) and (17) to a classical one-stage mathematical programming model. If we attempt to ascertain the "minimum" efficiency of specific DMU, we should use specific DMU lower bound outputs and other DMUs with lower bound inputs, and specific DMU upper bound inputs and other DMUs with upper bound outputs. Alternatively, we can find the specific decision making units to have the "maximum" efficiency, we should take the specific decision making unit's upper bound outputs and others' DMUs upper bound inputs, and the specific DMU lower bound inputs and others' DMU lower bound outputs. Then, the Model (16) and Model (17) can be transformed as follows:

$$\begin{aligned} &\text{Min} && (\rho_k)_\alpha^U = q - \frac{1}{m}\sum_{i=1}^{m} \frac{\left(S_i^-\right)^L}{(x_{ik})_\alpha^L} \\ &\text{Subject to} && q + \frac{1}{S}\sum_{r=1}^{S} \frac{\left(S_r^+\right)^U}{(y_{rk})_\alpha^L} = 1 \\ & && q(x_{ik})_\alpha^L = \sum_{\substack{j=1 \\ j \ne k}}^{n} (x_{ij})_\alpha^U \lambda_j^{'} + (x_{ik})_\alpha^L \lambda_k^{'} + \left(S_i^-\right)^L, i = 1, 2, \ldots, m \\ & && q(y_{rk})_\alpha^L = \sum_{\substack{j=1 \\ j \ne k}}^{n} (y_{rj})_\alpha^L \lambda_j^{'} + (y_{rk})_\alpha^U \lambda_k^{'} - \left(S_r^+\right)^U, r = 1, 2, \ldots, s \end{aligned} \tag{18}$$

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j^{'} = q, q > 0 \\ &\lambda_j^{'} \ge 0, j = 1, 2, \ldots, n \\ &\left(S_i^-\right)^L \ge 0, \left(S_r^+\right)^U \ge 0 \end{aligned}$$

$$
\begin{aligned}
&\text{Min} && (\rho_k)_\alpha^L = q - \frac{1}{m}\sum_{i=1}^{m}\frac{\left(S_i^-\right)^U}{(x_{ik})_\alpha^U} \\
&\text{Subject to} && q + \frac{1}{S}\sum_{r=1}^{S}\frac{\left(S_r^+\right)^L}{(y_{rk})_\alpha^L} = 1 \\
&q(x_{ik})_\alpha^U = \sum_{\substack{j=1\\ j\neq k}}^{n} (x_{ij})_\alpha^L\lambda_j' + (x_{ik})_\alpha^U\lambda_k' + \left(S_i^-\right)^U, i = 1, 2, \ldots, m \\
&q(y_{rk})_\alpha^L = \sum_{\substack{j=1\\ j\neq k}}^{n} (y_{rj})_\alpha^U\lambda_j' + (y_{rk})_\alpha^L\lambda_k' - \left(S_r^+\right)^L, r = 1, 2, \ldots, s \\
&&& \sum_{j=1}^{n}\lambda_j' = q, q > 0 \\
&&& \lambda_j' \geq 0, j = 1, 2, \ldots, n \\
&&& \left(S_i^-\right)^U \geq 0, \left(S_r^+\right)^L \geq 0
\end{aligned}
\tag{19}
$$

The optimal value $(\rho_k)_\alpha^L$ and $(\rho_k)_\alpha^U$ of Model (18) and Model (19) formulate DMU_k's interval efficiency $(\rho_k)_\alpha = \left[(\rho_k)_\alpha^L, (\rho_k)_\alpha^U\right]$ on given α–cut level.

According to the interval efficiency on given α–cut level, we can classify all the fuzzy decision making units $DMU_k (k = 1, 2, \ldots, n)$ into three classes:

(1) $E_\alpha^+ = \left\{DMU_k | (\rho_k)_\alpha^L = 1, k = 1, 2, \ldots, n\right\}$, if $DMU_k \in E_\alpha^+$, we say that it is fuzzy DEA efficiency on α–cut level;
(2) $E_\alpha = \left\{DMU_k | (\rho_k)_\alpha^L < 1, (\rho_k)_\alpha^U = 1, k = 1, 2, \ldots, n\right\}$, if $DMU_k \in E_\alpha$, we call that DMU_k is fuzzy DEA efficiency on α–cut level;
(3) $E_\alpha^- = \left\{DMU_k | (\rho_k)_\alpha^U = 1, k = 1, 2, \ldots, n\right\}$, if $DMU_k \in E_\alpha^-$, we say that it is fuzzy DEA inefficiency on α–cut level.

2.4 Super-SBM Model and Its Solution in Fuzzy Condition

Data Envelopment Analysis (DEA) mainly use various input and output of decision making units to construct efficient productive frontier, so we can measure the same kind decision making units' relative efficiency, and this is a nonparametric method. When using classical DEA models, we may find that many decision making units' efficiency are 1, thus, we can't have an efficiency order on them. In order to solve this problem, Andersen and Ptersen proposed Super-efficiency DEA in [12]. The difference between this model and conventional DEA model (CCR Model and BCC Model) is: classical DEA model, which put the linear combination of all decision making units as reference set of model evaluation, while Super-efficiency DEA model's reference set removes unvalued decision making units, consisting by the linear combination of other decision making units. So, when there are two or more than two decision making units (the efficiency value is 1), the measurement of inefficient units' efficiency is consistent between Super-efficiency DEA model and classical DEA model. While for the efficient decision making units, their efficiency can exceed 1, so the efficiency is different, thus we can have a complete order for all the decision making units.

2.4.1 Super-SBM Model

Assuming a production possibility set $\mathrm{P}(x_0, y_0)$ which covers (X, Y) but excluding (x_0, y_0) and n DMU_s, each with m inputs and s outputs, for the kth decision unit, its efficiency rating model is:

$$\begin{aligned} &\text{Min} \quad \tau_k = \frac{\frac{1}{m}\sum_{i=1}^{m}\frac{\bar{x}_i}{x_{ik}}}{\frac{1}{s}\sum_{r=1}^{s}\frac{\bar{y}_r}{y_{rk}}} \\ &\text{Subject to } \bar{x}_i \geq \sum_{\substack{j=1 \\ j\neq k}}^{n} x_{ij}\lambda_j, i=1,2, \ldots m \\ &\bar{y}_r \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} y_{rj}\lambda_j, r=1,2, \ldots, s \\ &\sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j = 1, j=1,2, \ldots, n \\ &\bar{x}_i > x_{ij}, 0 \leq \bar{y}_r \leq y_{rk}, \lambda_j \geq 0 \end{aligned} \tag{20}$$

where $(\bar{x}, \bar{y}) \in \mathrm{P}\backslash(x_k, y_k) \cap \{\bar{x} \geq x_k, \bar{y} \leq y_k\}$. Since Model (20) offers fractional linear programming, it may produce an infinite number of solutions. Tone [6] solved the above problem through multiplying non-negative scalar variable q for the denominator and numerator, assuming that the objective function of the denominator is equal to 1, so Model (20) transforms a simple linear programming similar to Model (21):

$$\begin{aligned} &\text{Min} \quad \tilde{\tau}_k = \frac{1}{m}\sum_{i=1}^{m}\frac{\bar{x}_i'}{\tilde{x}_{ik}} \\ &\text{Subject to} \quad \frac{1}{s}\sum_{r=1}^{s}\frac{\bar{y}_r'}{\tilde{y}_{rk}} = 1 \\ &\bar{x}_i' \geq \sum_{\substack{j=1 \\ j\neq k}}^{n} \tilde{x}_{ij}\lambda_j', i=1,2, \ldots, m \\ &\bar{y}_r' \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} \tilde{y}_{rj}\lambda_j', r=1,2, \ldots, s \\ &\sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j' = q, j=1,2, \ldots, n \\ &\bar{x}_i' > \tilde{x}_{ij}, 0 \leq \bar{y}_r' \leq \tilde{y}_{rj}, \lambda_j' \geq 0 \end{aligned} \tag{21}$$

2.4.2 Super-SBM Model's Fuzzy Expand and Solution

In a DMU_s set, we assume input $\tilde{x}_{ij}$ and output $\tilde{y}_{rj}$ are approximate available, also, $\tilde{x}_{ij}, \tilde{y}_{rj}$ of the membership functions are defined as $\mu_{\tilde{x}_{ij}}, \mu_{\tilde{y}_{rj}}$. According to the process of fuzzy expanding and solution of SBM model, we can get the optimal efficiency interval of DMU_k based on α–cut level:

$$
\begin{aligned}
&(\tau_k)_\alpha^U = min \frac{1}{m} \sum_{i=1}^{m} \frac{\left(\bar{x}_i^{'}\right)^L}{(x_{ik})_\alpha^L} \\
&\text{Subject to} \quad \frac{1}{s} \sum_{r=1}^{s} \frac{\left(\bar{y}_r^{'}\right)^U}{(y_{rk})_\alpha^U} = 1 \\
&\left(\bar{x}_i^{'}\right)^L \geq \sum_{\substack{j=1 \\ j \neq k}}^{n} (x_{ik})_\alpha^L \lambda_j^{'}, i = 1, 2, \ldots, m \\
&\left(\bar{y}_r^{'}\right)^U \leq \sum_{j=1}^{n} (y_{rk})_\alpha^U \lambda_j^{'}, r = 1, 2, \ldots, s \\
&\sum_{\substack{j \neq k \\ j \neq k}}^{n} \lambda_j^{'} = q > 0, j = 1, 2, \ldots, n \\
&\left(\bar{x}_i^{'}\right)^L \geq q(x_{ik})_\alpha^L, 0 \leq \left(\bar{y}_r^{'}\right)^U \leq q(y_{rk})_\alpha^U, \lambda_j^{'} \geq 0
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
&(\tau_k)_\alpha^L = min \frac{1}{m} \sum_{i=1}^{m} \frac{\left(\bar{x}_i^{'}\right)^U}{(x_{ik})_\alpha^U} \\
&\text{Subject to} \quad \frac{1}{s} \sum_{r=1}^{s} \frac{\left(\bar{y}_r^{'}\right)^L}{(y_{rk})_\alpha^L} = 1 \\
&\left(\bar{x}_i^{'}\right)^U \geq \sum_{\substack{j=1 \\ j \neq k}}^{n} (x_{ik})_\alpha^U \lambda_j^{'}, i = 1, 2, \ldots, m \\
&\left(\bar{y}_r^{'}\right)^L \leq \sum_{j=1}^{n} (y_{rk})_\alpha^L \lambda_j^{'}, r = 1, 2, \ldots, s \\
&\sum_{\substack{j \neq k \\ j \neq k}}^{n} \lambda_j^{'} = q > 0, j = 1, 2, \ldots, n \\
&\left(\bar{x}_i^{'}\right)^U \geq q(x_{ik})_\alpha^U, 0 \leq \left(\bar{y}_r^{'}\right)^L \leq q(y_{rk})_\alpha^L, \lambda_j^{'} \geq 0
\end{aligned}
\tag{23}
$$

2.5 Ranking of Decision Making Units in Fuzzy Condition

It is difficult to determine the ranking, because every decision making units' interval efficiency with α–cut level are measured in fuzzy condition. So we use the area measurement method proposed by Chen and Klien [7] to rank the fuzzy number at unknown membership functions:

$$\mathrm{I}\left(\tilde{E}_k, \tilde{R}\right) = \lim_{m \to \infty} \left(\frac{\sum_{i=0}^{m} \left[(E_k)_{\alpha_i}^{U} - c \right]}{\sum_{i=0}^{m} \left[(E_k)_{\alpha_i}^{U} - c \right] - \sum_{i=0}^{m} \left[(E_k)_{\alpha_i}^{L} - d \right]} \right) \tag{24}$$

where $c = min_{i,k}\left\{ (E_k)_{\alpha_i} \right\}$, and $= max_{i,k}\left\{ (E_k)_{\alpha_i} \right\}, i = 0, 1, \ldots, m$. If $\mathrm{I}\left(\tilde{E}_k, \tilde{R}\right)$ is higher than the represented the *k*th DMU ranking is higher.

3 Sample Data Introduction

3.1 The Source of Banks' Input-Output Data

As a financing institution which provide various services, banking industry's input-output is far different from general industry. According to existing literatures, [13] indicated that there are many methods to define banks' input-output, such as "Intermediary Approach", "Asset Approach", "Value Added Approach", "Production Approach", and "User Cost Approach", and the first three methods are usually been applied.

This article chooses the input-output index by applying Intermediary Approach and Asset Approach. And from abroad literatures about index selecting, we define input variable as deposit, net value of fixed assets and operating expenses; and define output variable as net credit and pretax profit.

In order to maintain consistency of statistical caliber and availability of data, this paper selects 13 listed commercial banks as sample. Including 5 state-owned banks: Agricultural Bank of China, Industrial and Commercial Bank of China, China Construction Bank, Bank of China and Bank of Communications; and 8 joint-stock commercial banks: China Citic Bank, China Minsheng Bank, Hua Xia Bank, China Everbright Bank, China Merchants Bank, Soci é t é G é n é rale, Shanghai Pudong Development Bank, and Ping An Bank. Ping An Bank came out after Shenzhen Development Bank Co., Ltd. Merged with Ping An Bank Co, SZPA at 2012. This paper uses the name of Ping An Bank to replace Shenzhen Development Bank during 2008–2011. The sample period is from the year 2008–2012, and sample data come mainly from every bank's annual report during the year 2008–2012.

3.2 How to Select Risk Factor Variable

Credit risks are the main risk during banks' operating activities. Commercial banks always use the bad loan ratio as one of the most important index measuring the credit risk. And it reflects the risk safety degree of banks' credit capital. The non-performing loan of commercial banks mainly includes substandard loan, doubtful loan and loss loan. The non-performing loan ratio = year-end non-performing loan of commercial bank/total loan of commercial bank. The higher bad loan ratio, the bigger risk of credit capital comes to commercial banks, and the less chance that banks can call in this loan, relatively, the greater loss may bring to the bank, and the bank's operating cost become bigger, so it will influence profitability and efficiency of operation.

Capital adequacy ratio = the capitalization of the bank/the weighted total assets. Capital adequacy ratio reflects how much loss the bank can undertake only depending on its own assets and shows the bank's ability to resist the risk when the depositors' and creditors' assets are impaired.

Past due loans are used to depict that repayment of bank loans are not on time, and reflect the service condition about this loans. Though past due loans can't transform to bad loan, we couldn't make sure if it will be get back, so it will cause uncertain loss to banks. And that continuous increase of past due loans would make the market have a negative attitude to the quality of banks' credit assets. Therefore, managing the past due loans not only can solve banks' past due loans quickly and reduce the uncertain loss, but also can promote the market's confidence to banks' credit security.

So, non-performing loan ratio, capital adequacy ratio and past due loans ratio are chosen as risk factor index to adjust Chinese commercial banks' input-output variables, then we get the efficiency by applying fuzzy DEA model.

This paper selects the research of the efficiency which has been adjusted by risk at 2011, so, we add the data of commercial banks' non-performing loan ratio, capital adequacy ratio and past due loans ratio which collected in 2011 to it based on former sample data base.

3.3 Adjust the Risk of Input-Output Variable

This article can not only measure 13 commercial banks' various efficiency in 2011 by applying BCC model and DEAP software, but also can get the difference about the input data. At the same time, we substitute the data which gained from every bank's non-performing loan ratio and capital adequacy ratio in 2011 into SFA model. Then, we can get the conclude by using frontier software. The results list in Table 1.

From Table 1, we can know that the non-performing loan ratio has a positive influence on the three input difference, and they are all significant at 0.01 level. It

Table 1 Parameter estimation of commercial banks' various efficiency in 2011

Dependent variable	The difference between the deposit			The difference between net value of fixed assets		
	Coefficient	Standard deviation	T value	Coefficient	Standard deviation	T value
Constant term	1333.300	1.892	704.749	33.794	1.904	17.747
The non-performing loan ratio	1007.002	9.567	105.262	22.970	1.176	19.526
Capital adequacy ratio	−189.592	2.023	−93.726	−4.427	0.197	−22.452
δ^2	185486.850			41.255		
Dependent variable	The difference between operating expenses					
	Coefficient	Standard deviation	T value			
Constant term	83.866	16.892	4.965			
The non-performing loan ratio	59.256	12.612	4.698			
Capital adequacy ratio	−11.510	2.128	−5.410			
δ^2	342.568					

indicates that the rising of non-performing loan ratio will aggravate the input difference, and the increase of the difference means the expansion of banks' input redundancy. Because of that, banks will be at a disadvantage operating condition and its efficiency will turn down. Nevertheless, capital adequacy ratio has a negative influence on the three input difference, and they are all significant at 0.01 level. The rising of capital adequacy ratio will decrease the input difference and banks' input redundancy. Because of that, banks' operating condition has improved and its efficiency also enhances. All of these are in line with theory and reality. Then, we can use Model (3), Model (4) and Model (5) to calculate triangle vague number based on risk adjustment. Moreover, the method to adjust output risk is to fuzzify net credit, and the vague number expression is $\tilde{L} = [L - L_0, L + L_0]$, among it, L is net credit, while L_0 is past due loan.

4 Empirical Analysis

4.1 Empirical Result Analysis After Risk Adjustment

First, put the variable without risk adjustment into SBM model and Super-SBM model to calculate the efficiency of 13 commercial banks. Then, put the input-output variable which has been adjust by risk into SBM model and Super-SBM model, after

Table 2 13 commercial banks' technical efficiency within risk adjustment in 2011

Bank	Model							
	SBM	Super SBM	Fuzzy SBM					
				$\alpha = 0$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 1$
Agricultural Bank of China	0.6234	0.6234	L	0.4772	0.5208	0.5512	0.5793	0.6234
			U	1.0000	1.0000	0.8736	0.7366	0.6234
Industrial and commercial Bank of China	1.0000	1.2213	L	1.2181	1.2191	1.2197	1.2204	1.2213
			U	1.2245	1.2236	1.2229	1.2223	1.2213
China Construction Bank	1.0000	1.0000	L	0.9236	0.9283	0.9338	0.9361	1.0000
			U	1.0008	1.0006	1.0004	1.0003	1.0000
Bank of China	1.0000	1.0409	L	1.0392	1.0396	1.0400	1.0403	1.0409
			U	1.0415	1.0413	1.0412	1.0411	1.0409
Bank of Communications	1.0000	1.0445	L	1.0343	1.0399	1.0410	1.0422	1.0445
			U	1.2031	1.1176	1.0824	1.0554	1.0554
China Citic Bank	1.0000	1.0113	L	1.0069	1.0079	1.0088	1.0100	1.0113
			U	1.0516	1.0396	1.0291	1.0169	1.0113
China Minsheng Bank	0.8418	0.8418	L	0.1950	0.2879	0.3747	0.4967	0.8418
			U	1.0000	1.0000	1.0000	1.0000	0.8418
Hua Xia Bank	0.6641	0.6641	L	0.0748	0.1381	0.1903	0.2706	0.6641
			U	1.0000	1.0000	1.0000	1.0000	0.6641
China Everbright Bank	0.8139	0.8139	L	0.1266	0.2053	0.2789	0.3886	0.8139
			U	1.0000	1.0000	1.0000	1.0000	0.8139
China Merchants Bank	0.8909	0.8909	L	0.2676	0.3709	0.4619	0.5815	0.8909
			U	1.0000	1.0000	1.0000	1.0000	0.8909
Soci é t é G é n é rale	1.0000	1.2032	L	1.0533	1.0847	1.1181	1.1496	1.2032
			U	1.2036	1.2035	1.2034	1.2034	1.2032
Shanghai Pudong Development Bank	1.0000	1.0270	L	1.0110	1.0140	1.0172	1.0226	1.0270
			U	1.0278	1.0277	1.0275	1.0273	1.0270
Ping An Bank	1.0000	1.4837	L	1.0147	1.0408	1.0711	1.1267	1.4837
			U	1.4931	1.4890	1.4865	1.4851	1.4837

that, we can get the 13 commercial banks' technical efficiency based on different α–cut set level within risk adjustment by applying DEA-Solver. On Table 2.

The second row in Table 2 gives out the 13 commercial banks' technical efficiency by applying SBM model in 2011. There are 5 commercial banks' efficiency less than 1, including Agricultural Bank of China, China Minsheng Bank, Hua Xia Bank, China Everbright Bank and China Merchants Bank, and they all are inefficiency unite of SBM. The other 8 commercial banks are efficiency unites of SBM, and they are all at efficiency frontier, meanwhile, there are neither input redundancy nor output deficiency. In order to rank them better, the third row lists out the 8 commercial banks' super efficiency by using Super-SBM model. From fifth row to ninth row, it depicts 5 commercial banks' technical efficiency by using Fuzzy SBM model and 8 commercial banks' technical efficiency by using Fuzzy Super-SBM model based on different α–cut set level. There are 5 levels about α–cut set, and it's a = 0 $\alpha = 0, 0.3, 0.5, 0.7, 1$. We can use α–cut set level as different degree of the risk adjustment about input-output, and the lower α–cut set level, the greater the degree of risk adjustment is. We also can get that the efficiency change interval is wider from the table. This means that the risk bank faced is bigger, and the variation of its efficiency is more intensive. For example, when $\alpha = 0$, it means that all of the banks need get their input-output risk changed in a greatest degree, and the degree of the corresponding efficiency is maximum. On the contrary, when $\alpha = 1$, it means that we don't do any risk adjustment to input-output, while the variables stay the same, at the same time, the upper limit value and lower limiting value of efficiency are equal to the efficiency calculated by SBM model and Super-SBM model.

From interval efficiency score in Table 2, we can get that risk factor has different influence on 13 commercial banks, and as the α–cut set level increasing, from $\alpha = 0$ to $\alpha = 1$, the amount of variation of commercial banks' efficiency is smaller. It means that the degree of banks' operating efficiency affected by the risk is lower as the risk which banks encounter become smaller, and the change is stable.

Among the state owned commercial banks, only Agricultural Bank of China is relative inefficiency, and the efficiency value is 0.6234. Meanwhile, the degree of its efficiency affected by the risk is greatest, because the efficiency interval is [0.4772, 1.000] when $\alpha = 0$. Amplitude of variation is bigger than other four state owned commercial banks at the same α–cut set level. When $\alpha = 0.5$, The upper limit efficiency value of Agricultural Bank of China has declined to 0.8736, so it is the weakest ability to resist risks among the state-owned commercial banks, and its risk-adjusted efficiency is lowest. From the Super-SBM model in Table 2, we can see that the efficiency of Industrial and Commercial Bank of China is highest, then it's Bank of Communications, following is Bank of China and China Construction Bank. From the degree affected by risk, Bank of China is minimally affected by risk, and the fluctuation range of efficiency is 0.023(=1.0415−1.0392) when $\alpha = 0$. The amount efficiency of variation of Industrial and Commercial Bank of China, China Construction Bank and Bank of Communications respectively are 0.064, 0.0772 and 0.1688. Above all, we know that 5 state-owned commercial banks' efficiency affected by risk ranking from weakest to greatest is Bank of China, Industrial and Commercial Bank of China, China Construction Bank, Bank of Communications and Agricultural Bank of China, while it also it risk adjustment order from highest to lowest of the 5 state-owned commercial banks.

Among joint-stock owned commercial banks, there are 4 banks which are SBM inefficiency, according to the efficiency order from highest to lowest, they orderly are China Merchants Bank, China Minsheng Bank, China Everbright Bank and Hua Xia Bank. There are 4 banks having efficiency, from Super-SBM model in Table 2, we can get that the efficiency of Ping An bank is highest, following is Soci é t é G é n é rale, Shanghai Pudong Development Bank and China Citic Bank. From the degree of risk impact, because joint-stock owned commercial banks start late, their risk management system is incomplete, and the management system is imperfect, the management level is not high, so the risk factor has a bigger influence on joint-stock owned commercial banks than state-owned commercial banks. Besides, the inefficient joint-stock owned commercial banks are greater affected by risk factor than state-owned commercial banks from Table 2. In the 4 inefficient commercial banks, Hua Xia Bank's efficiency amplitude fluctuation reach to 0.9252(=1.0000−0.0748) at α = 0, and it is greatest affected by risk factor, following is China Everbright Bank, China Minsheng Bank and China Merchants Bank, the efficiency amplitude fluctuation respectively are 0.8734, 0.8150 and 0.7342. Among the 4 efficient banks, Shanghai Pudong Development Bank's efficiency amplitude fluctuation reach to 0.0168(1.0278−1.0110) at α = 0, and it is weakest affected by risk factor, while Ping An Bank's fluctuation range of efficiency reach to 0.4784, and it is greatest affected by risk factor, the rest are Soci é t é G é n é rale (0.1503) and China Citic Bank (0.0447). Above all, we can rank the risk degree affecting joint-stock owned commercial banks' efficiency, from small to large, successively are Shanghai Pudong Development Bank, China Citic Bank, Soci é t é G é n é rale, Ping An Bank, China Merchants Bank, China Minsheng Bank, China Everbright Bank and Hua Xia Bank, so, Shanghai Pudong Development Bank's risk management level is highest, while Hua Xia Bank's risk management level is lowest.

4.2 *Commercial Banks' Efficiency Ordering After Risk Adjustment*

From Table 2, we can see that every commercial banks' efficiency based on α–cut set in the fuzzy condition. It is difficult to have a total ordering of all banks' efficiency affected by risk factor, so, using the interval evaluation method proposed by Chen and Klien [7] to rank the 13 commercial banks' fuzzy index, the result showing in Table 3.

Table 3 Commercial banks' efficiency ordering with risk adjustment in 2011

Bank	Order			
	SBM efficiency	Order	Fuzzy index I (*k*)	Order
Agricultural Bank of China	0.6234	13	0.3853	11
Industrial and Commercial Bank of China	1.2213	2	0.7825	2
China Construction Bank	1.0000	8	0.5770	8
Bank of China	1.0409	5	0.6360	5
Bank of Communications	1.0445	4	0.6535	4
China Citic Bank	1.0113	7	0.6168	7
China Minsheng Bank	0.8418	10	0.4034	10
Hua Xia Bank	0.6641	12	0.3557	13
China Everbright Bank	0.8139	11	0.3847	12
China Merchants Bank	0.8909	9	0.4249	9
Soci é t é G é n é rale	1.2032	3	0.7208	3
Shanghai Pudong Development Bank	1.0270	6	0.6208	6
Ping An Bank	1.4837	1	0.7831	1

Table 3 list the 13 commercial banks ordering according SBM efficiency and Fuzzy SBM interval efficiency, among that, interval efficiency is ordered by fuzzy index (k). According to the result, Ping An Bank's efficiency is the greatest among the 13 commercial banks, while the efficiency of Industrial and Commercial Bank of China is the highest among state-owned commercial banks, the gap between them is small. The efficiency of Agricultural Bank of China, Hua Xia Bank and China Everbright Bank have a different order after risk adjustment, because that SBM efficiency ranking doesn't count on risk factor, while interval efficiency ranking count in the risk factor. Agricultural Bank of China, as the state-owned bank, its risk management, risk system and risk control is better than Hua Xia Bank and China Everbright Bank, so we get this order.

5 Conclusion

This paper firstly uses input-output to have risk adjustment, then using Fuzzy SBM model and Super-SBM model to get Chinese 13 commercial banks' efficiency under risk condition based α–cut set interval efficiency in the year 2011. The results are as follows, first, the influence of risk factors on the 13 commercial banks in 2011 has varying degrees, and the less risk faced by banks, the smaller influence on the efficiency affected by the risk factors. Second, by comparing 13 commercial banks' efficiency fluctuation range of risk factors and find that among the state-owned commercial banks, Bank of China is the strongest ability to resist risks, Agricultural Bank of China is the weakest one; rather than the joint-stock commercial banks generally affected by a high degree of risk, among the joint-stock owned commercial banks, the risk management level of Shanghai Pudong Development Bank is highest, and Hua Xia Bank is the weakest one. Third, on 13 commercial banks sorting of risk-adjustment efficiency, we find that Soci é t é G é n é rale is the highest efficiency, Hua Xia Bank is the least efficiency, and the average ranking of the state-owned commercial bank is higher than the joint-stock commercial banks.

Acknowledgements Thanks to the support by The National Natural Science Foundation of China (No. 71201019) and Planning of Philosophy and Social Science Foundation of Liaoning (L13DJY065).

References

1. Sathye, M.: Efficiency of banks in a developing economy:the case of India[J]. Eur. J. Oper. Res. **148**(3), 662–671 (2003)
2. Hsiao, B., Chern, C.-C., Chiu, Y.-H., et al.: Using fuzzy super-efficiency slack-based measure data envelopment analysis to evaluate Taiwan's commercial bank efficiency[J]. Expert Syst. Appl. **38**(8), 9147–9156 (2011)

3. Eslami, R.,Khodabakhshi, M.,Jahanshahloo G.R.,et al.: Estimating most productive scale size with imprecise-chance constrained input-output orientation model in data envelopment analysis[J]. Comput. Ind. Eng. **63**(1), 254–261 (2012)
4. Jondrow, J., Materov, I., Lovell, C.A.K., et al.: On the estimation of technical inefficiency in the stochastic frontier production model. J. Econ. **19**, 233–238 (1982)
5. Tone, K.: A slacks-based measure of efficiency in data envelopment analysis. Eur. J. Oper. Res. **130**, 498–509 (2001)
6. Tone, K.: A Slacks-based measure of super-efficiency in data envelopment analysis. Eur. J. Oper. Res. (143), 32–41 (2002)
7. Chen, C.B., Klein, C.M.: A simple approach to ranking a group of aggregated fuzzy utilities. IEEE T. Syst. Man Cy. B. **27**, 26–35 (1997)
8. Huang, X., Yu, D., Yang, L.: A study on X efficiency of commercial banks in China–Based on three-stage DEA model[J]. Quant. Tech. Econ. **7**, 89–91 (2008)
9. Tang, Z.-Z.: Studing on risk and efficiency of commercial banks in China—Indexes choose based on factor analysis[J]. Mod. Bus. Tr. Ind. **6**, 162–164 (2009)
10. Hu, H., Song, Y.-L.: Analysis of operating efficiency on commercial banks based on affection of risk[J]. Econ. Res. Guide. **32**, 89–90 (2012)
11. van Laarhoven, P., Pedrycz, W.: A fuzzy extension of Saaty's priority theory. Fuzzy Set. Syst. **11**, 199–227 (1983)
12. Andersen, P., Petersen, N.C.: A procedure for ranking efficient units in data envelopment analysis. Manage. Sci. **39**(10), 1261–1264 (1993)
13. Bauer, P., Berger, A.N., Humphrey D.B.: Efficiency and productivity growth in US banking, In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S.(eds.) The Measurement of productive Efficiency: Techniques and Applications, pp. 386–413. Oxford University Press (1993)

A Novel Modeling and Analysis Approach to Efficiency Requirements for System of Systems

Qing-long Wang, Zhi-xue Wang, Yi-zhou Liu and Wei-xing Zhu

Abstract To solve the problem that the classical UML is lack of capability to model the fuzzy information for the C4ISR system, the paper proposes a novel modeling and analysis method based on fuzzy cloud model. It suggests a requirements analysis framework, which can define integrated requirements of both system functions and non-function, and extends the classic UML to improve the domain applicability. Then, it designs backward transformation algorithm of efficiency property and suggests a constructible method for efficiency evaluation function. Finally, the feasibility of the proposed approach is illustrated by using a military example.

Keywords Fuzzy UML · Efficiency requirements · Efficiency evaluation function · Cloud model

1 Introduction

Due to the rapid development of the society and industry, the system is the result of many, autonomous, heterogeneous, Constituent Systems (CSs), having complex constitution and interactions that generate large-scale complex System of Systems(SoS) [1]. The SoS development process contains the transform from capability requirements to high-level SoS requirements, and the satisfaction evaluating which the SoS performance meet the capability requirements. Among them, the key issues are modeling and automate verification process of SoS requirements [2].

The function requirements of SoS can be modeled by UML [3] or SysML as usual, but it is lacks of modeling method for efficiency requirements. The engineers usually need to translate the concept model of UML or SysML into an executable model,

Q. Wang · Z. Wang(✉) · Y.Liu · W. Zhu
College of Command Information, PLA University of Science and Technology, Nanjing 210007, China
e-mail: wzxcx801@163.com

Q. Wang
e-mail: jsxq910901@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_33

such as xUML [3], Petri Net [4]. However, it poses problems of model transformation between system model and executable model and their semantic equivalence. And there is a lot of fuzzy concept in the early stages of development [5].

Nowadays, for modeling C4ISR system (Command, Control, Communication, Computing, Intelligence, Surveillance and Resonance) which is typical SoS [6], it makes use of the architectural analysis method based on capability requirements. But it is hard to model the fuzzy information about the efficiency requirements of the C4ISR by the classic UML language. The fuzzy UML [7] as a UML profile can distinctly define the fuzzy information in the requirements.

In order to solve the modeling and analysis problems of C4ISR, we propose a method for fuzzy information of efficiency requirements, the aim of which is to help developers to build models for both function and fuzzy efficiency requirements. Therefore, it proposes the analysis framework of capability to acquire domain knowledge. Then, the fuzzy UML can be integrated in modeling SoS functional requirements and non-functional requirements. The efficiency evaluation of SoS can be evaluated quantitatively by using cloud model and its related algorithm. Finally, it realizes the synthetic modeling and analysis of SoS functional requirements and non-functional requirements by domain knowledge reuse technology.

2 Efficiency Requirements Modeling

2.1 Three-Layer Requirements Analysis Framework

Dong et al. propose three-layer requirements analysis framework [5], i.e. meta layer, domain layer and application layer. Requirements engineers acquire the domain knowledge from abstract concept at high level and domain specific concept, and build the requirements model with rich semantics. The requirements models are aimed to meet the needs of domain users. Based on the framework, this paper proposes a requirements analysis framework for SoS of C4ISR systems.

The analysis framework classifies the C4ISR requirements models into three layers: meta layer, domain layer and application layer. Specifically, the meta layer model is corresponding to the UML M2 layer model shown in Fig. 1. The meta model builds on the Capability Meta Model (CMM) which contains meta concepts and relations suggested by the Meta-Model Data Groups of Department of Defense Architecture Framework (DoDAF 2.0) [8]. This paper adds and fines some concepts, such as mission goal, capability, efficiency to build the meta model for the C4ISR system. The domain layer model, corresponding to UML M1 layer model, is the instantiation of the meta layer model. It is composed of domain concepts and relationship, such as the Warning capability or effectively intercept efficiency. Finally, the application layer is corresponding to UML M0 layer. The application model is at the bottom layer. They are the instantiations of domain concept and relationship model.

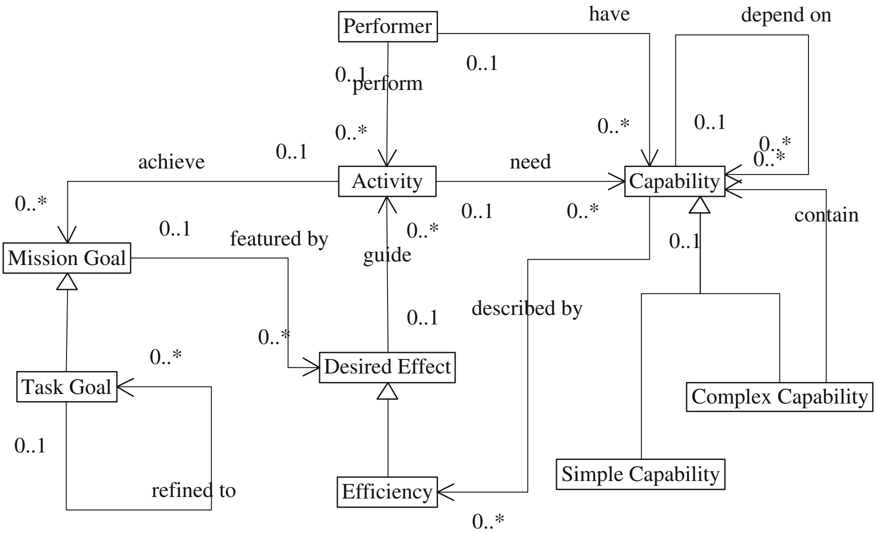


Fig. 1 The meta model of SoS requirements of C4ISR systems

2.2 *Three-Layer Requirements Analysis Framework*

In SoS, the efficiency requirements usually describe the implementation effect of system or capability. However, there are many vague and uncertain concepts that the classic UML is unable to model. In the domain knowledge of the capability requirements, the concepts become vague when add efficiency characteristics, for example, the early warning capability is a clear concept. But there is lack of clear and precise boundary for early warning, the early warning requirements, requiring as early as possible, become a vague concept that reflects the effect of capability.

In order to solve the modeling problems of vague domain knowledge, it introduces the fuzzy UML [7] in database modeling of complex objects. And it also facilitates modeling efficiency concepts of capability requirements by extended. The fuzzy class is defined of three kinds of fuzziness ($\mu : X \rightarrow [0, 1]$ is the membership function):

(1) A class is fuzzy, an object indeterminately belongs to a class. In the class properties, the membership attribution ($\mu \in [0, 1]$) is inserted to express the object membership;
(2) A property of class is fuzzy, namely, the range of property is a fuzzy class. And the fuzzy attribute is characterized as a keyword FUZZY to represent its vagueness;
(3) A class is fuzzy if its superclass or subclass is fuzzy class. It constructs the inheritance relationships between the fuzzy classes to describe this kind of fuzzy.

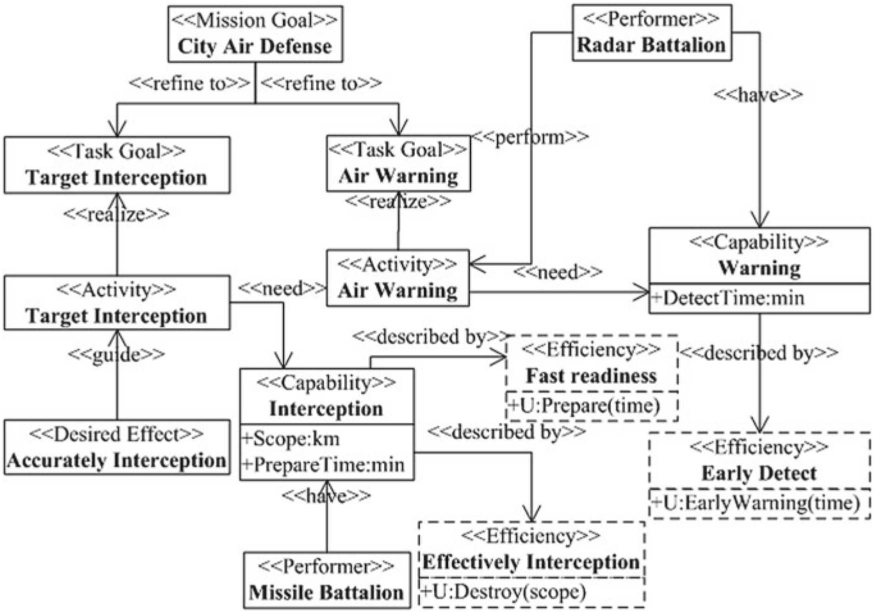


Fig. 2 The domain knowledge model of city air defense for C4ISR systems

This paper introduces the first kind of fuzziness to model the efficiency. A fuzzy concept can be modeled as a fuzzy class that has a membership function $\mu : X \rightarrow [0, 1]$. And it realizes the frame transform as a dashed box to express the fuzzy class, as shown in Fig. 2. The efficiency evaluation function is discussed distinctly to determine the objects membership degrees in Sect. 3.

In the domain of city air defense, the city air defense tasks require an early warning capability. Besides, the desired effect of early warning capability hopes for the earliest possible detection of air targets to provide the best warning effect. However, the early warning is a fuzzy and uncertain concept that can be modeled as a fuzzy class. The other fuzzy concepts, i.e. rapid operational readiness and effective target damaged, etc., also can be modeled as fuzzy class.

3 Efficiency Requirements Modeling

3.1 Efficiency Evaluation Based on Cloud Model

In the above, the efficiency values can be modeled as a fuzzy value μ of the first layer fuzzy. In order to realize the evaluation of efficiency value of the system, this section proposes the efficiency evaluation function to evaluate the efficiency value.

Dong et al. [5] defines the efficiency in the viewpoint of probability that the system can meet the desired effect in prescriptive operating conditions and time. The Efficiency Evaluation Function (EEF) is a calculation that converts the system efficiency characteristics to probability value of realized the desired effect.

Furthermore, it is quite important for modeling capability requirements to determine the form of the EEF. In analysis of efficiency requirements of C4ISR system, some attributes of performance or efficiency can be provided a quantitative description. The EEFs calculate those quantitative attributes to obtain the corresponding probability of achieving the desired effect. It is makes it great difficult to analyze system efficiency. But the cloud model proposed a means to solve the above problem.

3.1.1 Cloud Model

Cloud model is an uncertainty conversion model of qualitative and quantitative based on the traditional fuzzy sets and probability and statistics.

Definition 1 Let U be a universal set of discourse described by precise numbers, and C be the qualitative concept associated with U. If there is a number $x \in U$, which randomly realizes the concept C and the certainty degree of x for C, i.e. $\mu(x) \in [0, 1]$, is a random value with stable tendency $\mu : U \to [0, 1]$, $\forall x \in U$, $x \to \mu(x)$. Then the distribution of x on U is defined as cloud. Every x becomes a cloud drop [9]. A cloud model can be described with three numerical characteristics, i.e. $Ex(Expected\ value)$, $En(Entropy)\ and\ He(Hyper\text{—}Entropy)$, shown in Fig. 3. The numerical characteristics of merging fuzziness and randomness provide mathematical basis for mixed problem of qualitative and quantitative.

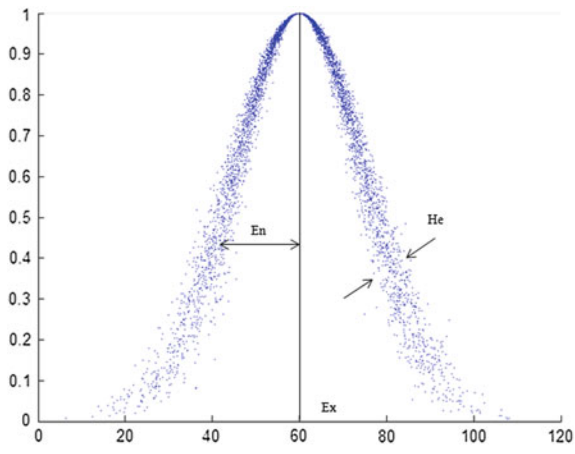


Fig. 3 The cloud model and its numerical characteristics

3.1.2 The Normal Cloud Model

Definition 2 Let U be a universal set of discourse described by precise numbers, and $C(Ex, En, He)$ be the qualitative concept associated with U. If there is a number $x \in U$, which randomly realizes the concept C, and x submit to Gaussian distribution $x \sim (Ex, (En')^2)$. Among them, En'^2 is a random value with Gaussian distribution $En' \sim N(En, He^2)$, and the certainty degree of x for C, i.e. $\mu(x) = e^{-\frac{(x-Ex)^2}{2(En')^2}}$. Then the distribution of x on U is defined as normal cloud [9].

For the efficiency evaluation of C4ISR SoS, it is convenient for normal cloud model to construct EEF to translate fuzzy qualitative information into accurate quantitative information, which can achieve a better efficiency evaluation of the system.

3.2 The Efficiency Evaluation Function Based on Cloud Model

3.2.1 The Outline of EEF Construction

Due to the practical situation C4ISR system for effectiveness evaluation, this section uses the cloud model theory to evaluate the efficiency of the system by testing data or expert experience. The EEF construct process is as follows:

(1) According to efficiency evaluation requirements for C4ISR system, the efficiency data is collected, and the collected data is preprocessed for the coming calculation.
(2) The system efficiency data are translated into several cloud models of EEF by the Backward Cloud Transformation of EEF.
(3) The cloud models of EEF are combined into one cloud model synthesis of EEF for C4ISR system.
(4) According to the cloud model synthesis of EEF, it is easy for us to solve the expectation function, that is, the EEF. The EEF can compute the synthesis efficiency value of C4ISR systems.

3.2.2 Design for Backward Cloud Transformation of EEF

The system gains many items of data and the data are corresponding to m set of data. If every unit of date is considered as a cloud droplet, m units of cloud will be gained by the reverse builder, namely the backward cloud transformation of EEF.

In the procedure of efficiency evaluation, the system has multiple cloud models of efficiency attribute which need to be combined into a synthesis efficiency cloud model. Since the importance of each cloud model is different, every models weight must be considered before the combination. In this section, it introduces the concept of a weighted percentage to reflect the importance of each cloud model.

The weighted percentage of cloud models is equal to the fraction that the corresponding number of data points has divided by the total number of data points, namely weighted percentage ω, i.e.,

$$\omega = \frac{\lambda_i}{N} \tag{1}$$

wherein, N is the total number of data points, and λ_i is the number of date point for each attribute owned by the cloud models. The construct process of backward cloud transformation is displayed in Algorithm 1

Algorithm 1 Backward cloud transformation algorithm of EEF

Output: The m units of attribute cloud $(EC_1, EC_2, \ldots, EC_m)$ and the numerical characteristics $(Ex_1, \ldots, Ex_m, En_1, \ldots, En_m, He_1, \ldots, He_m)$

Step 1 Calculate data sample mean of each group for $X_i(x_{i1}, x_{i2}, \ldots, x_{iM_i}), i = 1, 2, \ldots, m$.

$$\overline{X_i} = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{ij}. \tag{2}$$

Step 2 Calculate the expected value of each cloud model, i.e., $(Ex_1, Ex_2, \ldots, Ex_m) = (\overline{X_1}, \overline{X_2}, \ldots, \overline{X_m})$

Step 3 Calculate the entropy of each cloud model, i.e.,

$$En_i = \sqrt{\frac{1}{M_i - 1} \sum_{j=1}^{M_i} (x_{ij} - Ex_i)^2}. \tag{3}$$

Step 4 For each cloud droplet $(x_{ij}, \mu_i), j = 1, 2, \ldots, M_i$, it can be calculated as follows,

$$En'_{ij} = \sqrt{\frac{-(x_{ij} - Ex)^2}{2 \ln \mu_i}}. \tag{4}$$

Step 5 Calculate the Hyper Entropy He_i from the Standard Deviation of each En'_i, i.e.,

$$He_i = \sqrt{\frac{1}{M_i - 1} \sum_{j=1}^{M_i} (En'_{ij} - \overline{En'_i})^2}. \tag{5}$$

Therefore, the algorithm can calculate the numerical characteristics $EC_i\,(Ex, En, He)$ of each cloud of efficiency attributes.

3.2.3 The Synthetic Efficiency Evaluation Function

In order to evaluate the efficiency of system, the synthetic EEF can be calculated by the each weight of cloud of efficiency attribute, which is based on the each related efficiency attribute cloud generated from the sample data.

According to the numerical characteristics of each efficiency attribute cloud and corresponding weights, it calculates the new synthetic cloud for efficiency evaluation, i.e.,

$$\begin{cases} Ex = \sum_{i=1}^{m} (Ex_i \times \omega_i) \\ En = \sqrt{\sum_{i=1}^{m} (En_i^2 \times \omega_i)} \\ He = \sum_{i=1}^{m} (He_i \times \omega_i) \end{cases} \quad (6)$$

wherein, m is the sum of the efficiency attribute clouds, and $\omega_i(\sum_{i=1}^{m} \omega_i = 1)$ is the corresponding weights of each cloud of efficiency attribute. In the light of aforementioned formulas, it is convenient to obtain the mathematic characteristics of synthetic cloud of efficiency evaluation $EC_S(Ex, En, He)$. Then, its expectation function can be achieved by mathematic characteristics, i.e.,

$$\mu(x) = e^{-\frac{(x-Ex)^2}{2En^2}} \quad (7)$$

In the function curve shown in Fig. 5, each point $[x_i,\ \mu(x_i)]$ is corresponding to the expectation Ex_i of each cloud droplets. In the efficiency evaluation of C4ISR system, the expectation curves are the qualitative descriptions of some fuzzy concepts. Consequently, the curve equation is the EEF. The certainty degree of the cloud droplet is unrelated with mathematic characteristic of the concept and the connotation of the concept, which is the characteristic of cloud model. This goes to show that different experts with different acknowledge have the same cognitive rules on the same efficiency concept of the same C4ISR system. This also eliminates the biases of experts acknowledge and improve the scientificity and reliability of the efficiency evaluation result.

4 A Case Study

A hypothetical case is presented to illustrate the feasibility of the proposed modeling and analysis method for efficiency requirements of C4ISR system, which is based on a city defense system.

There is a hypothetical major city demand for protection of the urban area center and a certain radius of protected area. A missile battalion carries out the duty to provide the city protection, deploys into short-range and medium-range air defense missiles to counter the difference range for city protection. According to aforesaid three-layer requirements analysis framework, the application task concept of protection the X City task is a fuzzy and uncertain concept in application model, with evaluation of EEF (Importance(distance)) for urban important degree (the assumption is that if the city scale is larger, its importance is higher). According to the characteristics of city air defense mission, the domain experts firstly define a reasonable rule: if a city possesses an important degree of n ($0 < n < 1$), then the destroyed effect for deployed target interception capability must be greater than n to complete the mission. In addition, there is an extra assumption that if Importance(10) = 0.9 and Importance(35) = 0.7. According to the supposed scenario, it builds a C4ISR system model for city air defense, as shown in Fig. 4.

After modeling the system of city air defense, it demands to evaluate the systemic efficiency for further analysis whether the system can successfully complete the city air defense mission. Firstly, through the aforementioned cloud model-based constructing method of EEF, an EEF can be constructed from the sample data given by domain experts. Subsequently, the EEFs evaluate systemic efficiency for different missions and different operational environment. Finally, the results of evaluation provide the foundations for optimizing the system requirements and architecture.

In order to accomplish the mission of city air defense, the target destroy efficiency of air defense missile must be greater than the efficiency value of Importance attribute. The target destroy capability contain efficiency property of effective target destroy, which is evaluated by efficiency evaluation function, namely Destroy(scope)

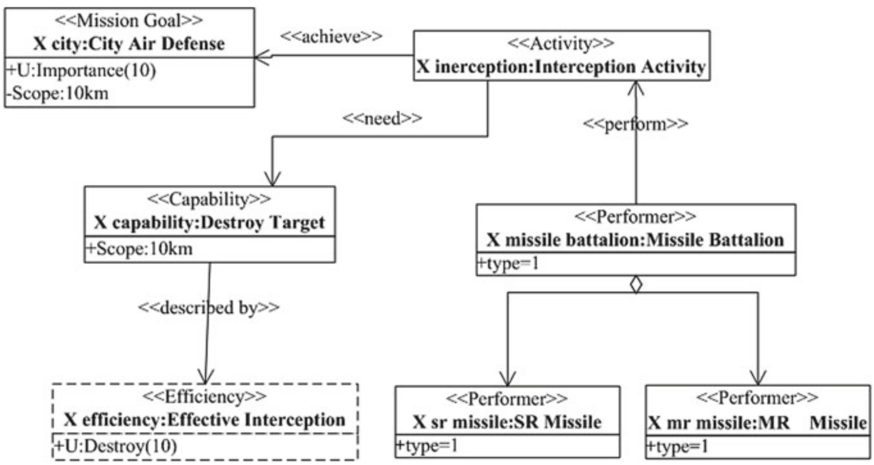


Fig. 4 The application model of C4ISR system for city air defense

(scope is variable of destroy range), as shown in Fig. 2. Moreover, the air defense missions are divided into medium-range and short-range air defense missions for responding different invading targets.

Therefore, it is feasible to make use of the proposed method for EEF construction based on cloud model. Firstly, it starts off with four groups of sample points provided by experts. Then four cloud model of efficiency attribute can be gotten by the proposed formula, i.e. $EC_1(10,\ 6,\ 0.5)$, $EC_2(12,\ 7,\ 0.5)$, $EC_3(14,\ 8,\ 0.5)$, $EC_4(16, 9, 0.5)$. The four cloud models of efficiency attribute combine into on synthetic cloud model and the EEF of the efficiency obtains in expectations curve of the synthetic cloud model. The mathematic characteristics of synthetic cloud model is $EC_S(Ex = 20.1,\ En = 4.85,\ He = 0.5)$ in this $\omega_1 = 0.15$, $\omega_2 = 0.2$, $\omega_3 = 0.25$, $\omega_4 = 0.25$. The synthetic cloud model is shown in the Fig. 5. The membership function is $\mu(x) = e^{-\frac{(x-20.1)^2}{2\times 4.85^2}}$.

The membership function is the EEF Destroy(scope). With observing the curve of cloud model, the widest destroy range is almost 25 km for short-range air defense missiles (the membership value is greater than 0.5), and the optimal destroy efficiency is nearly in the range of 20 km.

Since changes of the mission task cause changes in the sample data, the EEF should also change in order to accurately reflect the sample point. There is an assumption that the air defense missions transform from short-range missions to medium-range mission. Moreover, the domain experts provide the sample data points based on empirical data for medium-range mission, a total of four sets of sample data points. The four cloud models are obtained by the backward cloud transformation, i.e. $EC_1(30, 8, 0.5)$, $EC_2(35, 9, 0.5)$, $EC_3(40, 10, 0.5)$, $EC_4(45, 11, 0.5)$. And the synthetic cloud model is $EC_L(Ex = 37.5,\ En = 9.56,\ He = 0.5)$, in this $\omega_1 = 0.2$, $\omega_2 = 0.3$, $\omega_3 = 0.3$, $\omega_4 = 0.2$. Its evaluation function is $\mu(x) = e^{-\frac{(x-37.5)^2}{2\times 9.56^2}}$.

In Fig. 6, it compares the changes in EEF Destroy(scope) after the radius of air defense mission changes. It is distinct to observe the alteration of EEF Destroy(scope) in different missions corresponding change of sample data.

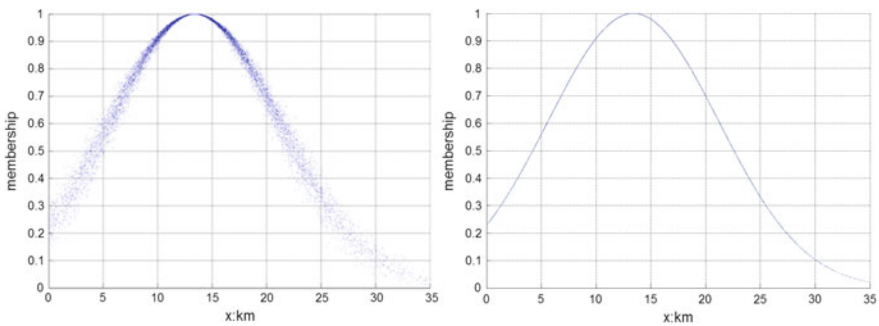


Fig. 5 Synthetic cloud model and EEF for short-range air defense missile

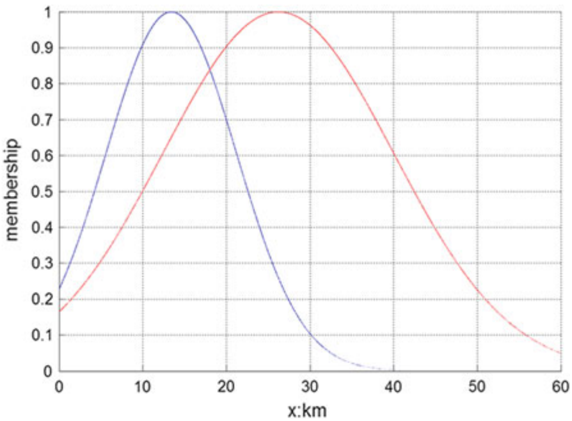


Fig. 6 The compare curve of the EEF Destroy(scope) for different missions

The missile intercepting range is arranged to 10 km in stipulated city protection scope of application model, and shown in Fig. 6. At this point, the destroy efficiency of X short-range missile is 0.9, the destroy efficiency of X medium-range missiles is 0.5, therefore the destroy efficiency of X medium-range missiles do not meet the requirements of the mission. Since the city protection scope and missile attack range is expanded to 35 km, the destroy efficiency is 0.05 for X short-range missile and is 0.8 for X medium-range missiles. The destroy efficiency of short-range missiles no longer meet the new task request. And it must use the medium-range missile to complete the new mission.

The model verification results of two kinds of mission embody the advantages of the method. On the one hand, the method may check out defective non-function requirements that cannot meet the mission needs in RE process. On the other hand, the method provides reference materials for technology and military staffs selecting the weapon equipments of different mission context.

5 Conclusion

This work proposes a fuzzy extension-based modeling and analysis method for efficiency requirements to solve the modeling and analysis problems of non-function requirements in capability requirements analysis for C4ISR system. A three-layer requirements analysis framework is built for the comprehensively function and non-function requirements model. Moreover, a fuzzy UML extension facilitates the description of fuzzy and uncertain concepts in the analysis of efficiency requirements and extends domain applicable scope. And the EEF constructing approach based on cloud model address quantitative description problems in systemic efficiency evaluation, which performs the EEF to achieve the systemic efficiency evaluation by quantitative description. The results of systemic efficiency evaluation facilitate optimizing system architecture.

Acknowledgments Thanks to the support by National Natural Science Foundation of China (No. 61273210) and the Meta-Model Data Groups of Department of Defense Architecture Framework for the contribution to the meta-models.

References

1. Ncube, C., Lim, S., Dogan, H.: Identifying top challenges for international research on requirements engineering for systems of systems engineering. RE 2013, 342–344 (2013)
2. Ender, T., Leurck, R., Weaver, B.: System of systems analysis of ballistic missile defense architecture effectiveness through surrogate modeling and simulation. IEEE Syst. J. **4**(2), 156–166 (2010)
3. Haimes, Y.: Modeling complex systems of systems with phantom system models. Syst. Eng. **15**(3), 333–346 (2012)
4. Ge, B., Hipel, K., Yang K., Chen, Y.: A novel executable modeling approach for system of systems architecture. IEEE Syst. J. **8**(1), 4–13 (2014)
5. Dong, Q., Wang, Z., Chen, G.: Domain-specific modeling and verification for C4ISR capability requirements. J. Central South University **19**(5) 1334–1341 (2012)
6. Petersen, K., Khurum, M., Angelis, L.: Reasons for bottlenecks in very large-scale system of systems development. Inf. Softw. Technol. **56**(10), 1403–1420 (2014)
7. Ma, Z., Zhang, F., Yan, L., Cheng, J.: Representing and reasoning on fuzzy UML models: a description logic approach. Expert Syst. Appl. **38**(3), 2536–2549 (2011)
8. US Department of Defense. DoD architecture framework version 2.0 (Volume I-II-III) [S/OL], MOD Partner (2009). http://www.us.army.mi/suite/page/454707
9. Wang, G., Xu, C., Li, D.: Generic normal cloud model. Inf. Sci. **280**, 1–15 (2014)

Exploring High Dimension Large Data Correlation Analysis with Mutual Information and Application

Yu-shan Jiang, Dong-Kai Zhang, Xiao-min Wang and Wen-yan Zhu

Abstract Applying for information entropy theory, we present a measure of dependence for multi-variables relationships: the high dimensional maximal mutual information coefficient (*HMIC*). It is a kind of maximal information-based nonparametric exploration (*MINE*) statistics for identifying and classifying relationships in large data sets which generalizes the maximum information coefficient (*MIC*) measurement in mutual variables. To decreasing the complexity of the *HMIC* computing, the improved uniform grid is proposed by data grid idea. At the same time, some optimal single axis partition algorithm (*SAR*) is built to ensure the feasible of the *HMIC* measurement. Finally we apply the *HMIC* to analysis the data sets of physical measurement among college students.

Keywords Entropy · Maximal mutual information coefficient · Single axis partition

1 Introduction

People from different fields analyze a variety of datasets to understand human behaviors, find new trends in society, and possibly formulate adequate policies in response. Typically, we address the problem of finding interesting and unknown patterns via data mining methodology. Data mining enables people to extract information from a data set and convert it into a comprehensible structure for further use. The

Y. Jiang (✉)
State Key Laboratory of Integrated Automation of Process Industries,
Northeastern University, Shenyang 110004, China
e-mail: jys@neuq.edu.cn

Y. Jiang · X. Wang · W. Zhu
School of Mathematics and Statistics,
Northeastern University, Qinhuangdao 066004, China

D.-K. Zhang
School of Mathematics and Information Science,
Shijiazhuang University, Shijiazhuang 050035, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_34

nonparametric methods, such as the χ^2 tests, the Fisher exact probability test, and the Spearman rank correlation, have long been among the standard tools of the statisticians. Recently, some new nonparametric or 'distribution-free' statistical methods [1–12], have gained prominent in statisticians. In [13] The kernel estimators was used for control engineering dynamical systems as an objective function of automatic control. [7] discusses some multiple comparisons nonparametric approaches with the computation of adjusted p-value which improve the results offered by the Friedman test in some circumstances. Using Maximum Likelihood [9] analysis the molecular sequence data which improves the computational efficiency and the accuracy of the estimates. [11] performed a systematic analysis of interaction dynamics across different technologies and show the high-throughput yeast two-hybrid is the only available technology for detecting transient interactions on a large scale.

Recently Reshef [1] detected a novel associations in large data sets which called the maximal information coefficient (*MIC*) measure of dependence for two-variable relationships. They have proved that the *MIC* of a noiseless functional relationship converges to 1 as sample size grows. And if the sample distribution (X, Y) are statistically independent the *MIC* converges to zero. Some applications on *MIC* to data sets in global health, gene expression and major-league baseball are also studied by them. Intuitively, *MIC* is based on the idea that if a relationship exists between two variables, then a grid can be drawn on the scatter plot of the two variables that partitions the data to encapsulate that relationship. In [2] the generalized 3*MIC* algorithm has been built with quite low algorithm complexity $O(n^2)$. However, the control parameter $B(n)$ is inevitably determined with experience. Therefore some uniform grid G is proposed to improve this defect in this study. The uniform grid G is independent with the data size. Furthermore we establish *SAR* recurrence algorithm. By proof we show its feasible. Finally we demonstrate the application of 3*MIC* to students physical measurement.

The organization of the study is as follows. In Sect. 2 we present some high dimensional mutual information which generalized the common mutual information and establish the definition of *HMIC* in high statistical distributions $(X_1, \ldots, X_m)$. In Sect. 3 we present the single axis partition algorithm (*SAR*) to approach the *HMIC* in three dimension variables distributions (X, Y, Z) based on the group theory and the separability of the entropy. Finally, in Sect. 4 the simulation application example shows that the recurrence algorithm is efficiency in practice.

2 Main Definitions on High Dimensional Mutual Information Coefficient

Given a finite set $D = \{(x_1, \ldots, x_m) | x \in \mathbb{X}_1, \ldots, \mathbb{X}_m\}$ of m dimensions, To simplify the computational complexity we partition D into $\alpha_1 \times \cdots \times \alpha_m$ uniform grid G which we call $G(\alpha_1 \ldots \alpha_m)$. There are some definitions and lemmas that we will use in the following section.

Definition 1 For given grid $G(\alpha_1 \ldots \alpha_m)$, let $D|_G$ be the distribution induced by the data points in D on the grid $G(\alpha_1 \ldots \alpha_m)$. That is

$$D|_{G(\alpha_1 \ldots \alpha_m)} = \left\{ p(\xi_1, \ldots, \xi_m) | p(\xi_1, \ldots, \xi_m) = \frac{N(\xi_1, \ldots, \xi_m)}{|D|} \right\}$$

where $1 \le \xi_i \le \alpha_i (1 \le i \le m)$, $N(\xi_1, \ldots, \xi_m)$ is the number falling in the box $g(\xi_1, \ldots, \xi_m)$ of G, $|D|$ is the size of data set D. For given data set D and grids G one can obtain the distributions $D|_G$. Instead of original data we focused on the distributions $D|_G$. For fixed $D|_G$ make grid family as the follows

$$F_G = \{D|_{G'} : D|_{G'} \subset D|_G\}$$

where if the grid G^1 can be further split into the grid G^2 then the relation between D_{G^1} and D_{G^2} is defined as $D_{G^1} \subset D_{G^2}$. We make the following definition of *HMIC*.

Definition 2 For a finite set $D \subset \mathbb{R}^m$ with data size $|D| = L$ and positive integers $\alpha_1, \ldots, \alpha_m$, the high dimensional maximal mutual information coefficient (*HMIC*) of a set D under the grid $G(\alpha_1, \ldots, \alpha_m)$ is given by

$$HMIC(D, \alpha_1, \ldots, \alpha_m) = \max_{D|_{G'} \in F_G} \{C(D|_{G'})\} \quad (1)$$

where $C(D|_{G'})$ denotes the characteristic matrix.

$$C(D|_G) = \frac{I(D|_{G'})}{\min\{\log(\alpha_1), \ldots, \log(\alpha_m)\}} \quad (2)$$

$I(D|_{G'})$ is the m-variables mutual information deduced by distribution grid $D|_{G'}$.

$$I(X_1; \ldots; X_m) = \sum_{i=1}^{m} H(X_i) - \sum_{\substack{1 \le i,j \le m \\ i \ne j}} H(X_i, X_j) + \cdots + (-1)^{m-1} H(X_1, \ldots, X_m) \quad (3)$$

$H(X)$ is the entropy of X which is defined by

$$H(X) = -E[\log p(x)] = -\sum_{x \in \mathbb{X}} p(x) \log p(x)$$

$H(X_{i_1}, \ldots, X_{i_r})$ is the corresponding joint entropy of $X_{i_1}, \ldots, X_{i_r}$. Since for any given grid $G'(\alpha_1, \ldots, \alpha_m)$

$$-\min\{\log \alpha_1, \ldots, \log \alpha_m\} \le I(D|'_G) \le \min\{\log \alpha_1, \ldots, \log \alpha_m\}$$

therefore all the characteristic matrix fall between −1 and 1. The symmetry of *HMIC* implies that the characteristic matrix $C(D|_G)$ remains the same when the axis of D are interchanged. In particular, we use fixed uniform grid G instead of $B(\varepsilon)$ in [1] for control parameters practicality.

3 *SAR* Algorithm for Generating the 3*MIC*

In this section we describe our algorithm for generating the *HMIC* of a set of three variables data(3*MIC*). By Definition 2 the core of generating 3*MIC* is finding an optimal sub grid distribution D_G^* in F_G which return the highest three variable mutual information where the three variable mutual information is defined by

$$I(X_1;X_2;X_3) = \sum_{i=1}^{3} H(X_i) - \sum_{\substack{1\le i,j\le 3\\ i\ne j}} H(X_i,X_j) + H(X_1,X_2,X_3).$$

To this end we propose some axis sort method on the elements of F_G with group theory firstly.

Definition 3 (*Partial order relation of* F_G) The order '$\prec$' corresponding to set family F_G is defined as

$$D_{G^1} \prec D_{G^2} \Leftrightarrow G^1 \subset G^2, D_{G^1}, D_{G^2} \in F_G.$$

Above definition will be used in our following algorithm.

3.1 F_G Representation and Description of the Optimal Function

For some given uniform grid $G(m,n,h)$ in three dimension space (X,Y,Z) we rewrite it as

$$G(m,n,h) = (P,Q,R)$$

where $P = P[x_0,x_1,\dots,x_m]$ be the x-axis partition and $x_i (i = 1,\dots,m)$ is the grid nodes, $x_0 = \inf_{x\in X}\{x\}$, $x_m = \sup_{x\in X}\{x\}$. $Q[y_0,y_1,\dots,y_n]$ and $R[z_0,z_1,\dots,z_h]$ are the y-axis and the z-axis partitions respectively. Correspondingly the order '$\prec$' with axis partition is defined as

$$P_1[x_0,\dots,x_l] \prec P_2[x_0,\dots,x_h] \Leftrightarrow \{x_0,\dots,x_l\} \subset \{x_0,\dots,x_h\}.$$

Now we give the definition of grid family F_G.

Definition 4 The grid family of $G(m, n, h)$ is defined as

$$F_G = \{G' = (P', Q', R') | P' \prec P, Q' \prec Q, R' \prec R\}.$$

For example, if the uniform grid $G = G(1, 2, 3)$ with axis partitions $P = [0, 1]$, $Q = [0, 1, 2]$ and $R = [0, 1, 2, 3]$ then the grid family of F_G can be represented by

$$F_G = \left\{ \begin{array}{cc} \{[0,1],[0,2],[0,3]\}, & \{[0,1],[0,1,2],[0,3]\}, \\ \{[0,1],[0,2],[0,1,3]\}, & \{[0,1],[0,1,2],[0,1,3]\}, \\ \{[0,1],[0,2],[0,2,3]\}, & \{[0,1],[0,1,2],[0,2,3]\}, \\ \{[0,1],[0,2],[0,1,2,3]\}, & \{[0,1],[0,1,2],[0,1,2,3]\} \end{array} \right\}.$$

The $3MIC$ is the maximum $C(D|_G)$ of the given data set D induced from above grid family. We begin by outlining an idealized algorithm for generating the characteristic matrix. Algorithm 1 represents what we would use if efficiency were not a problem.

Algorithm 1 The Outline of $3MIC(D, G)$

Require: D is a set data with three variables
1:**for** $G'(m, n, h)$ such that $G' \in F_G$ **do**
2: $G'(m, n, h)$ grid on D
3: $C(D|'_G) \leftarrow \frac{I(D|'_G)}{\min\{\log(m), \log(n), \log(h)\}}$
4:$3MIC_G \leftarrow \max\{C(D|'_G)\}$
5:**end for**
6:**return** $3MIC_G$

However the rank of F_G grows exponentially with the size $(m \times n \times h)$ of the uniform grid G. Thus, considering the symmetry property of $3MIC$, we build some single axis recurrence algorithm (SAR) in the next subsection.

3.2 *Single Axis Recurrence Algorithm Approach*

Assume that our set $\{(x, y, z)|(x, y, z) \in D\}$ is uniform grided by $G(m, n, h)$. We denote the uniform partitions of the axis by specifying the nodes of the exist interval of the data. For the case of x-axis the interval is $[\inf_{(x,y,z)\in D}\{x\}, \sup_{(x,y,z)\in D}\{x\}]$ and the uniform partition is

$$B^x = [b_0^x(\inf\{x\}), b_1^x, \dots, b_m^x(\sup\{x\})].$$

All subgrid $G' \subset G$ are $P = [p_0^x, \dots, p_s^x]$ with $p_0^x < p_1^x < \dots < p_s^x$ and $P \prec B^x$. The corresponding probability distribution is

$$D_P = \{p_i | p_i = \frac{sum\ of\{p_{i-1}^x \le x < p_i^x\}}{|D|}\}.$$

Similarly

$$D_Q = \{q_i | q_i = \frac{sum\ of \{p^y_{i-1} \leq y < p^y_i\}}{|D|}\},$$

and

$$D_R = \{r_i | r_i = \frac{sum\ of \{p^z_{i-1} \leq z < p^z_i\}}{|D|}\}.$$

are the y, z axis distributions respectively. Now fixed y, z-axis partitions Q, R, using our definition on mutual information in three variables, we have $I(X; Y; Z) = H(X) - H(X|Z) + H(Z|X, Y) - H(Z|Y)$ where $H(.)$ denotes Shannon entropy. However, since Y, Z are fixed, the SAR algorithm need only maximize $H(X) - H(X|Z) + H(Z|X, Y)$ over all partitions P on x-axis of D in order to maximize $I(P; Q; R)$. Thus, we show Proposition 1 to approach the SAR.

Proposition 1 *For given uniform grid $G(m, n, h)$ with y, z-axis partitions Q, R. And a data set D of size N. For every $1 < m' \leq m$, define*

$$I^*(m', Q, R) = \max_{|P'| \leq m'} \{H(P') - H(P'|R) + H(R|P, Q)\} \tag{4}$$

where P' is all partitions of size up to m' under the y, z resolution of $n' \times h'$. Define N_i is the number of points in the i-th layer of G and $N_{i,j}$ to be the number of points in the (i, j) column of G and $N_{i,j,k}$ to be the number of points in the (i, j, k) box of G. we have the following recurrence for $m' > 1$.

$$I^*(m', Q, R) = \max_{1 \leq i < m} \big\{I^*(m' - 1, Q, R) + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|P')$$
$$- H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|R) - H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|Q) + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|Q, R)\big\}$$

where $\widehat{N}_i$ is the number of points in the layer the new node $\widehat{x}_i$ added in. $\widehat{x}_i^-, \widehat{x}_i^+$ are the adjacent left and right nodes of $\widehat{x}_i$.

Proof For given x axis partition $P'_{m'}$ of size m' let $P'_{m'-1}$ be the $m' - 1$ partition such that $P'_{m'-1} \prec P'_{m'}$. Considering $I'(P'_{m'}; Q; R) = H(P'_{m'}) - H(P'_{m'}|R) + H(R|P'_{m'}, Q)$ we have

$$\begin{aligned}
&I'(P'_{m'}; Q; R) \\
&= \sum_{i=1}^{m'} \frac{N_i}{N} \log \frac{N}{N_i} - \sum_{k=1}^{h'} \frac{N_k}{N} \sum_{i=1}^{m'} \frac{N_{i,k}}{N_k} \log \frac{N_k}{N_{i,k}} + \sum_{i=1}^{m} \sum_{j=1}^{n'} \frac{N_{i,j}}{N} \sum_{k=1}^{h'} \frac{N_{i,j,k}}{N_{i,j}} \log \frac{N_{i,j}}{N_{i,j,k}} \qquad (5)\\
&= I'(P'_{m'-1}, Q, R) + \frac{\widehat{N}_i^+}{N} \log \frac{\widehat{N}_i}{\widehat{N}_i^+} + \frac{\widehat{N}_i^-}{N} \log \frac{\widehat{N}_i}{\widehat{N}_i^-} - \sum_{k=1}^{h'} \frac{N_k}{N} \left(\frac{\widehat{N}_{i,k}^+}{N_k} \log \frac{\widehat{N}_{i,k}}{\widehat{N}_{i,k}^+} + \frac{\widehat{N}_{i,k}^-}{N_k} \log \frac{\widehat{N}_{i,k}}{\widehat{N}_{i,k}^-} \right) \\
&\quad + \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^+}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{\widehat{N}_{i,j}^+} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j,k}^+} + \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^-}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{\widehat{N}_{i,j}^-} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j,k}^-} - \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j}} \log \frac{\widehat{N}_{i,j}}{\widehat{N}_{i,j,k}}
\end{aligned}$$

where $\widehat{N}_i$ is divided into $\widehat{N}_i^+, \widehat{N}_i^-$, i.e. $\widehat{N}_i = \widehat{N}_i^+ + \widehat{N}_i^-$. Noticing that $N_{i,j} = \sum_{k=1}^{h'} N_{i,j,k}$ and $N_i = \sum_{j=1}^{n'} N_{i,j}$. By the the additivity property of information entropy [14] it follows from (5) that

$$\begin{aligned}
I'(P'_{m'};Q;R) = {} & I'(P'_{m'-1};Q;R) \\
& + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|P') - H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|R) \\
& + \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^+}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{\widehat{N}_{i,j}^+} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j,k}^+} + \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^-}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{\widehat{N}_{i,j}^-} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j,k}^-} \\
& - \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}}{N} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j}} \log \frac{\widehat{N}_{i,j}}{\widehat{N}_{i,j,k}} \\
= {} & I'(P'_{m'-1};Q;R) + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|P') - H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|R) \\
& + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{N} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j,k}^+} + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{N} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j,k}^-} \\
& - \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{N} \log \frac{\widehat{N}_{i,j}}{\widehat{N}_{i,j,k}} - \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{N} \log \frac{\widehat{N}_{i,j}}{\widehat{N}_{i,j,k}} \\
= {} & I'(P'_{m'-1};Q;R) + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|P') - H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|R) \\
& + T_1 + T_2 - T_3 - T_4.
\end{aligned}$$

For the last four term of the above equation $T_1 + T_2 - T_3 - T_4$, we have

$$\begin{aligned}
T_1 + T_2 - T_3 - T_4 &= (T_1 - T_3) + (T_2 - T_4) \\
&= \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{N} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j,k}^+} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j}} + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{N} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j,k}^-} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j}} \\
&= \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{N} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j}} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j,k}^+} + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{N} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j}} \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j,k}^-} \\
&= \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^+}{N} \log \frac{\widehat{N}_{i,j}^+}{\widehat{N}_{i,j}} + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^+}{N} \log \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j,k}^+} \\
&\quad + \sum_{j=1}^{n'} \frac{\widehat{N}_{i,j}^-}{N} \log \frac{\widehat{N}_{i,j}^-}{\widehat{N}_{i,j}} + \sum_{j=1}^{n'} \sum_{k=1}^{h'} \frac{\widehat{N}_{i,j,k}^-}{N} \log \frac{\widehat{N}_{i,j,k}}{\widehat{N}_{i,j,k}^-} \\
&= -H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|Q) + H([\widehat{x}_i^-, \widehat{x}_i, \widehat{x}_i^+]|Q, R).
\end{aligned}$$

Taking the maximum of both sides of the equation we have the desired result.

If given some number x of partitions and some number y, z of partitions, we could run the SAR algorithm function on every possible y, z size partitions, we would find an optimal grid. But the number of possible y, z partitions makes this infeasible. A natural approach to this problem is to consider only grids for which at least one axis is partitioned ahead. To this end, the SAR algorithm fixes an partition of y, z axis with Q, R and then runs to the result. Later, the SAR is called again but with the axes switched. The maximum of the three scores obtained is used.

4 Analysis of Physical Measurement in College Students

Among thousands of college student's data with dozens of body index considering the efficiency one need to find several most important relations to represent the student's physical condition. To solve the problem we obtained the colleges data sets about the students' physical measurement of different grade in recent five years which are listed in Table 1. For each college in each grade the data sets properties include gender (G), height (H), weight (W), boss mass index (BMI), lung capacity (LC), lung capacity weight index (LWI), endurance scores (ES), flexibility scores (FS) and smart scores (SS).We removed all students who's BMI were not in the normal range ([60, 100]). Figure 1 shows the visual relationship among lung capacity, flexibility scores and smart scores data. The data sets was analyzed using the 3*MIC* algorithm. Table 2 shows the highest and the lowest relationships computing by 3*MIC*. In Figs. 2 and 3 the upper part three surfaces with contour are derived from the characteristic matrices of the data sets D. The uniform grid is $G = G(40, 40, 40)$. In each surface x-axis represents the fixed y, z uniform grid number used to partition the last two variables of D, y-axis represents the grid number being to partition the first variable. The z-axis represents the characteristic value under those grids

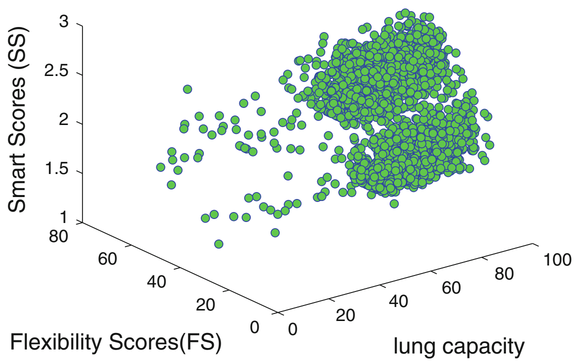


Fig. 1 Visualizations of three variables Spatially. The relationship among LC, FS and SS is not evident from the figure

Table 1 Pretreatment on physical measurement variables

G	H (cm)	W (kg)	LC (ml)	ES	FS	SS
M	181.8	62.7	4980	78	44.8	2.27
M	176.9	65.4	3980	81	38.1	2.68
M	181.5	74.1	4321	81	38.7	2.3
FM	167.2	64.9	3500	94	27.9	1.74
…	…	…	…	…	…	…

Table 2 3*MIC* scores on physical measurement variables

Physical variables			3MIC scores
ES	FS	SS	0.0046
…	…	…	…
H	W	LC	0.28

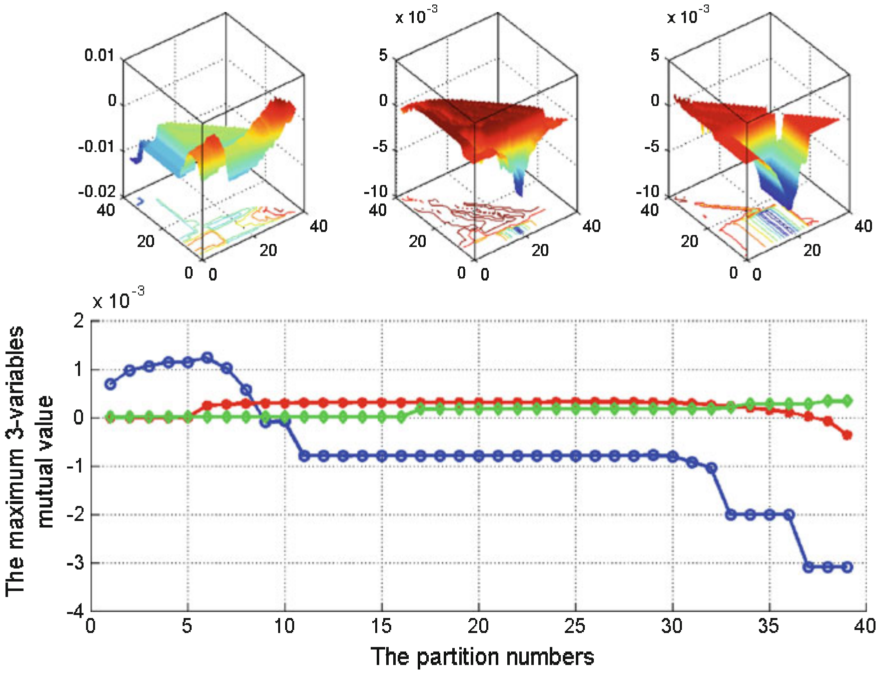


Fig. 2 Visualizations of the characteristic matrices and the 3MIC with the data set $D = \{(ES, FS, SS)\}$

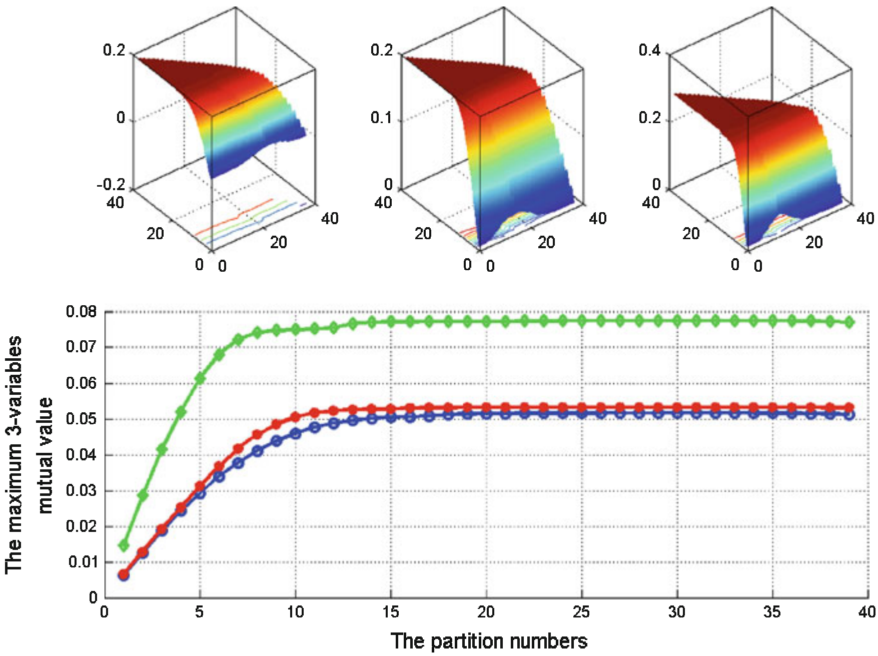


Fig. 3 Visualizations of the characteristic matrices and the 3MIC with the data set $D = \{(H, W, LC)\}$

$G'(m', 40, 40)(1 < m' \leq 40)$. The sample size is 2514. On bottom figure is the 3MIC curve in different grids. It is consistent with the experience.

5 Conclusion

This study describes the improved *HMIC* algorithm that finds coefficients in multi-variables $(X_1, \ldots, X_m)$. We have showed that the *HMIC* Algorithm is different from the original one in complexity and properties. The application domain of the proposed algorithm is not limited to physical data. In contrast with the maximum grid numbers the proposed uniform grid $D|_G$ and $F|_G$ can automatically determined. As we known the order of $F|_G$ is the important factor affected program running speed. To avoid the dimensional disaster induced by $F|_G$, the *SAR* recurrence process was proposed for three variables maximal mutual information coefficient. This *SAR* process is a dynamic optimization process. And the local optimal result I^* is also the global optimal. On the other hand we need give a complete proof of the convergence of the algorithm theoretically. This is our after time work.

Acknowledgments The research is supported by NNSF of China under Grant No. 61273008 and No. 61104003. and the Fund of Hebei Education Department No.z2014096. The research is also supported by the Key Laboratory of Integrated Automation of Process Industry (Northeastern University). The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that allow us to improve the manuscript.

References

1. David, N.: Reshef: detecting novel associations in large data sets. Science **334**, 1518–1524 (2011)
2. Jiang, Y., Zhang, Q., Liu, C.: Some novel measurement on exploring large data sets based on multil-variables mutual information theory. J. Theor. Appl. Inform. Technol. **2**(47), 547–550 (2013)
3. Zeshui, X.: Intuitionistic fuzzy aggregation operators. IEEE Trans. Fuzzy Syst. **15**(6), 1179–1187 (2007)
4. Karpinets, T.V., Park, B.H., Uberbacher, E.C.: Analyzing large biological datasets with association networks. Nucl. Acids Res. **40**(17), 1–8 (2012)
5. Wang, L., Wang, X.: On the worst case data sets for order statistics. Appl. Math. Inf. Sci. **2**(6), 356–362 (2012)
6. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M.: Wim Hordijk and Olivier Gascuel: new algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. **3**(59), 307–321 (2010)
7. García, S., Fernández, A., Luengo, J., Herrera, F.: Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. Inf. Sci. **10**(180), 2044–2064 (2010)
8. Erceg-Hurn, D.M., Mirosevich, V.M.: Modern robust statistical methods. Am. Psychol. Assoc. **7**(63), 591–601 (2008)
9. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S.: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol. **10**(28), 2731–2739 (2011)
10. Albanese, D., Filosi, M., Visintainer, R.: cmine, minerva and minepy: a C engine for the MINE suite and its R and Python wrappers, pp. 1–10 (2012). arXiv:1208.4271 [stat.ML]
11. Das, J., Mohammed, J., Haiyuan, Y.: Genome-scale analysis of interaction dynamics reveals organization of biological networks. Bioinformatics **28**(14), 1873–1878 (2012)
12. Deng, X., Havukkala, I., Deng, X.: Large-scale genomic 2D visualization reveals extensive CG-AT skew correlation in bird genomes. BMC Evol. Biol. **234**(7), 1471–1482 (2007)
13. Kulczycki, P.: Nonparametric estimation for control engineering. In: 4th WSEAS/IASME International Conference on Dynamical Systems and Control, pp. 115–121 (2008)
14. Cover T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)

Strong Robustness Control of Multi-modal Based on Human-Simulated Intelligence for Uncertainty System

Bo Bi and Hu Xu

Abstract Uncertainty process is the puzzle in the actual control engineering, and aimed at being difficult to actualize the effective control by conventional PID controller, the paper explored a sort of control strategy of strong robustness based on multi-modal control algorithm. In the paper, it made the anatomy of cybernetics characteristic for uncertainty system, pointed out the limitation of application in PID controller, researched on dynamic and static characteristics human simulated intelligent controller, and based on the multi-modal control, constructed the control algorithm of a uncertainty system. Under the strong pulse interference, it took a two-order with time lag process control as an example, made the contrastive experimental simulation respectively by PID and multi-modal control algorithm based on HSIC, and the process response demonstrated that the HSIC control algorithm owned very strong robustness control performance. The result of experimental research shows that it is feasible and reasonable to uncertainty process by proposed multi-modal control algorithm based on HSIC.

Keywords Uncertainty process · Multi-modal control algorithm · Robustness performance · Human simulated intelligent control

1 Introduction

The PID and its improved controller have a very wide range of applications in the industrial automation control system, and so far, it still dominates in the application in the automation control field. The advantage is that it is easy to use and adjust, and clear in physical concept, but there are limitations in their application, and generally

B. Bi (✉)
International School, Chongqing Jiaotong University, Chongqing 400074, China
e-mail: 804805129@qq.com

H. Xu
College of Automation, Chongqing University, Chongqing 400044, China
e-mail: 1597725908@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_35

it is limited to the control of deterministic process. If the controlled object is not a deterministic process, when the process state deviates from a range of adjustment, then it is very difficult to make the correction of state departure by retuning parameters. Because the PID control is a kind of quantitative control paradigm, the control is based on strict mathematical model. For the uncertainty process, the PID controller is incapable of action because of its difficult to mathematical modeling. The control method that can avoid the mathematical modeling is that the intelligent controller would be adopted, for example, fuzzy logic control, artificial neural network control, expert control system and the human simulated intelligent control etc. The significant feature adopting intelligent controller is that as long as the controlled process deviates from the expected process state, the controller will adopt various means to actualize the control over the process, and makes it return to the expected goal state. Among them, the most typical is the human simulated intelligent controller (HSIC), it can adopt the control model combined the qualitative control with quantitative decision without strict mathematical modeling, and it is suitable for many kinds of complex process control with uncertainty. The control of correlating the complex process uncertainty is currently facing the control puzzle [1], and the following are some necessary exploration on relevant issues.

2 Facing Challenges of Uncertainty System Control

2.1 Characteristics of Process Cybernetics

The cybernetics characteristic has determined that it is difficult to build the mathematics model for uncertainty process, and it reflects in many aspects [2]. ① the controlled complex object is not high in structured degree, it demonstrates the characteristic of semi-structure and non-structure, and therefore, it is difficult to make an accurate mathematical description. ② There is the correlated coupling between variables that reflect the process state in the system, and the relation among them sometimes is uncertain, and thus, it is difficult to achieve the decoupling among the variables so as to realize the simple control of single variable. ③ there is serious nonlinear characteristics in the controlled object, and therefore, it is difficult to make the quantization processing because of uncertainty effect. ④ there is the characteristic of dispersity, time-varying, unknown and random in system parameters, and it owns the uncertainty in system parameter and time-varying delay. ⑤ Due to system environment being complex, the interference of the external environment is often unknown, diversity and random, and therefore it has a striking feature of uncertainty. In view of being difficult to make the mathematical modeling, in general it can only be described by generalized knowledge model, and so it is very difficult to control for complex correlation system with uncertainty.

2.2 Facing Challenges of Process Control

The conventional control is based on the certainty description of mathematical model, and the uncertainty results in being difficult to build the mathematical model. The model parameter and the structure changes in a large range, some industrial interference is difficult to predict, its control process is the morbid structure, and it leads to be difficult to achieve the expected control by conventional control. The input and output of the conventional control is difficult to make the information interaction with the outside world, it does not accept the information data of non-quantity form, and only the intelligent control can be as the input and output by the information data of non-quantitative form. The task of conventional control is single, and the output is either constant (constant value control), or following the expected trajectory (track control). But the task of complex control system is always complex, it often requires the ability of automatic planning and decision-making etc., and the conventional control can not meet the control requirements. For the linear control problem, the conventional control theory and mature technology can be used for reference, but for the control problem with uncertainty, so far, it is still lack of effective means of settlement. The conventional control does not own the ability to fuse relevant knowledge such as the controlled object, environment and human being control strategy etc., and its application scope is limited to simple control system. The conventional control is difficult to show the characteristics of hybrid control process and generalized knowledge model by mathematical description form. The modal of the conventional control mode is difficult to embed other control strategies, and its control modal is relatively simple. For example, it can not adopt the control modal combined qualitative decision with quantitative control as well as the multi-modal control modal combined open-loop with closed-loop control mode. The conventional control structure is fixed, and it has not the ability of variable structure and self-organization etc. Of course, but it does not own the ability of self-compensation, self-repair and decision-making etc.

Facing the above challenges, by using the intelligent control strategy it can obtain more satisfactory solutions. Because it can be convenient to introduce the human control experience, control expert knowledge, wisdom and skills of the site operator and so on. The controller is designed based on the generalized control model, and combined mathematical model with knowledge system it uses the integrated control strategy fused the human intelligence. Obviously the conventional control is incapable of action.

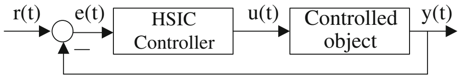


Fig. 1 Generalized control model of HSIC

3 Control Strategy of Human Simulated Intelligence

3.1 Control Model

Human intelligence can be reflected from the ability aspects of observation, learning, understanding and cognition about things. In fact, the human intelligence is the ability to understand and adapt to all kinds of behavior (including the ability to control behavior). Intelligent control is a kind of automatic control technique, and it is emphasized on the prerequisite that it can autonomously drive the intelligent controller to achieve the expected control goal without manual intervention in the actualizing control process. The control strategy of HSIC is reflected in the aspects of simulation for human intelligence, the control method is intuitive, and it finds the answer to solve the problem directly from the human being itself. When simulating the control behavior of human being, it can commence from two aspects of control function and control structure. Based on artificial intelligence and automatic control theory, it can summarize up the knowledge and control experience of the actual operators and control experts about controlled object, and describe the inspiration and intuition reasoning as well as control behavior by using the production rule in artificial intelligence. Based on features memory and online characteristics identification, it can abstract out the dynamic characteristics model of system, and actualize the multimodal control based on HSIC by means of combination among qualitative decision, quantitative control and open-closed loop control. In the aspects of functional structure, it is different in information processing and decision making mechanism of HSIC at different hierarchy. In view of system error and its change rate being measurable, on these grounds it can carry out on-line feature recognition and feature memory for control system, and consequently, it can be used to construct the control model and control algorithm based on HSIC. The generalized control model of HSIC is shown as in Fig. 1, in which, r(t) represents the input of control system, e(t) is the system error, u(t) shows the output of HSIC controller and y(t) is called as the output of control system.

3.2 Static Characteristics of HSIC

In the HSIC control, the sketch map of static characteristics is shown as in Fig. 2.

In Fig. 2, the motion trajectory of OA-AB-BC-CD-DE-EF-FH-HG-GI shows the transfer process of the system state in the motion space, and at the same time, it also displays and explains that which kind of control mode should be taken in the process of adjusting control. For example, in OA section, the proportional control

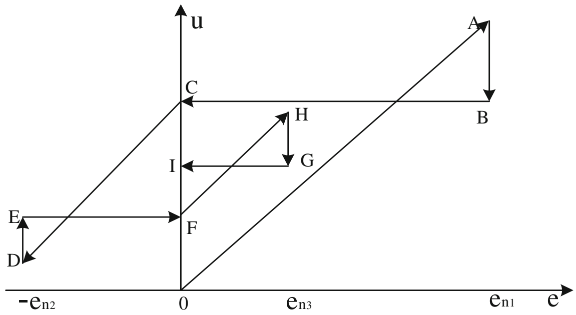


Fig. 2 Static characteristics of HSIC

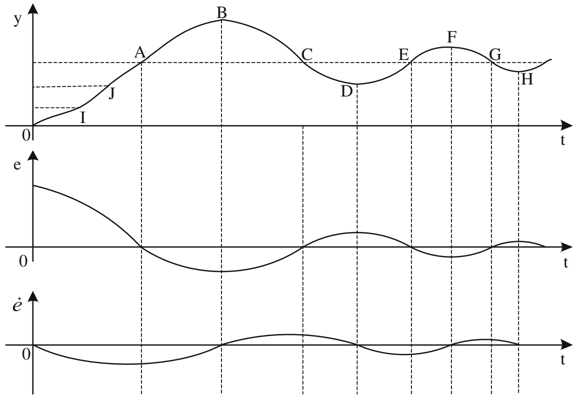
mode should be taken, and the output u = K_p e. In which, K_p is a proportional coefficient. The feature is $e\dot{e} > 0$, after error e reached the first peak value e_{n1}, it enters the inhibition control mode of AB, and so, the operating range is over interval [0, e_{n1}]. The feature of AB segment is to make the original proportional coefficient K_P be changed as K_P k, in which, $k \not> 1$, namely AB is the inhibition control section, and when the motion trajectory reaches to point B, the output would be reduced down to $u_{01} = k\ K_p\ e_{n1}$, and then it enters the BC mode. The feature of BC section is that the controller output keeps unchanged, but the e value is decreased from extremum to the origin of the 0. The analysis of subsequent motion trajectory of CD-DE-EF is similar to the motion trajectory of OA-AB-BC, but the action direction of the control cycle is opposite with previous control cycle because of the error e being negative. Similarly, it can analyze the motion trajectory change of FH-HG-GI in third control cycle. After multiple cycles of control, the system can automatically achieve the expected steady state.

3.3 Dynamic Characteristics of HSIC

The dynamic characteristics of general system can be represented by output response, and a typical dynamic response characteristic of HSIC is shown as in Fig. 3.

From the analysis of response curve of 0A-AB-BC-CD-DE-EF-FG-GH in Fig. 3, it can be seen that there is the relation between system error e and its change rate $\dot{e}$ in different output time piece. Through analyzing relation among e, $\dot{e}$ and y, it can find to deal with the control strategy, and according to the different feature mode of system error, it can adopt different control strategy. For example, in section of OA, AB, BC, CD, DE, the feature of each section respectively is $(e > 0, \dot{e} < 0)$, $(e < 0, \dot{e} < 0)$, $(e < 0, \dot{e} > 0)$, $(e > 0, \dot{e} > 0)$, $(e > 0, \dot{e} < 0)$, and so on., and aimed at different situation, it can adopt different control mode. If $e \cdot \dot{e} > 0 \cup e = 0 \cap \dot{e} \neq 0$, then it can takes the proportional control, and the control amount u changes with the error according to the proportion. If $e \cdot \dot{e} < 0 \cup \dot{e} = 0$,then it takes keeping control mode, and the control amount u keeps the cumulative sum of error e maximum. If

Fig. 3 Dynamic characteristics of HSIC



$e \cdot \dot{e} > 0 \cup e = 0 \cap \dot{e} \neq 0$, then it can takes the proportional control, and the control amount u changes with the error according to the proportion. If $e \cdot \dot{e} < 0 \cup \dot{e} = 0$, it takes the keeping control mode, and the previous keeping value should be greater than the cumulative sum of error e maximum. If $e \cdot \dot{e} > 0 \cup e = 0 \cap \dot{e} \neq 0$,it still takes proportional control, and so on. After through repeated proportion-keeping control mode, the final control u converges to a constant value, while the system error e converges to zero.

From above analysis in static and dynamic characteristics of HSIC, it can be seen that the basic control algorithm of HSIC is to simulate the thinking process of human being, and its control essence is that aimed at different characteristics of the system error, it selects different control mode and control algorithm [3, 4].

4 Control Algorithms: Basic Algorithm and Fusion Algorithm

4.1 Basic Control Algorithm

Based on the analysis of static and dynamic characteristics of HSIC [5], according to the basic feature of combining open-loop control with closed-loop control, the strategy of HSIC can be induced as two kinds of system dynamic error basic modes. ① If $e \cdot \dot{e} \leq 0$ or $e = \dot{e} = 0$, then it takes the semi-open-loop keeping control node. ② If $e \cdot \dot{e} \geq 0$ and $e + \dot{e} \neq 0$, then it takes the proportional control mode. The basic control algorithm of HSIC is as the following.

$$U = \begin{cases} K_p \cdot e + k \cdot K_p \cdot \sum_{i=1}^{n-1} e_{m,i} & (e \cdot \dot{e} > 0 \cup e = 0 \, \dot{e} \neq 0) \\ k \cdot K_p \cdot \sum_{i=1}^{n} e_{m,i} & (e \cdot \dot{e} < 0 \cup \dot{e} = 0) \end{cases}$$

In the algorithm, $e_{m,j}$ is the jth peak value of error maximum, U is the controller output, and the meaning of other mathematical symbol is the same as previous mentioned above. The feature of basic control algorithm is to realize the double modal control, namely for different control mode it takes different control strategy, and taken the form of open-loop control and closed-loop control alternately, it realizes the control for controlled object.

4.2 Control Algorithm of Intelligence Fusion

In order to improve the control quality of controller and simplify the structure of control system, it can fuse the control expert knowledge, control experience of human being, wisdom and skills of the site operator, specific control rules, inference engine and knowledge base etc. as a body, and form a fusion intelligent controller based on human simulated intelligence. The outstanding advantage of fusion control model is that it can be convenient to adopt the production rules "If <condition> Then <action>" so as to establish the knowledge set and control rule base. Because there is no direct link between the rules, it is good in modularity and naturality. According to different conditions, each piece of control rule can be independently made the addition and modification, and therefore, it owns very good ability to adapt the change of environment. If $e \cdot \dot{e} > 0 \cup e = 0\, \dot{e} \neq 0$ then it takes proportional control mode, and if $e \cdot \dot{e} < 0 \cup \dot{e} \neq 0$ then it takes keeping control mode. In addition, by means of the two-order derivative characteristic of system error, the controller performance can be further improved. After fused the adjusting skill of human being, practice experience, expert knowledge, operating skill and wisdom and so on, the special control algorithm can be induced by structured English description method. For example, aimed at the uncertainty system of complex correlation, one of the methods can be constructed as the following.

If $e \cdot \dot{e} \geq 0$ and $\neq 0$ Then

If $\dot{e} \cdot \ddot{e} \geq 0$ Then

$P(t) = \bar{P}_{n-1} + K_p e + P_{HD}(\text{Note}, P_{HD} = \tilde{P}_{l-1} + kK_p \dot{e})$

If $\dot{e} \cdot \ddot{e} < 0$ Then

$$P(t) = \tilde{P}_{n-1} + K_p e + P_{HD}(\text{Note}, P_{HD} = kK_p \sum_{i=1}^{l} \dot{e}_{m,i})$$

If $(e \cdot \dot{e} < 0\ \text{OR} |e| + |\dot{e}| \leq \delta)$ Then

If $|e| \geq \frac{1}{2}|e_{m,n}| > \delta$ Then

$P(t) = \bar{P}_n + kK_p(e - \frac{1}{2}e_{m,n})$

Else $P(t) = \bar{P}_n$

In which, $\bar{P}_n = kK_P \sum_{i=1}^{n} e_{m,i}$, P is the output of HSIC controller, defines $\bar{P}_0 = 0$ (memory of error extremum), PHD is the semi-derivative component of proportional control mode, and $\bar{P}_n$ is the needed keeping value of the n th output. Defines $\tilde{P}_0 = 0$ (derivative extremum memory), $\ddot{e}$ is the two-order derivative with time of error, $\dot{e}_{m,i}$ is the i th extremum of $\dot{e}$ in the proportional control mode, $\tilde{P}_l$, is the l th needed keeping constant value of semi-derivative output component P_{HD} in proportional control mode. i, l, n is respectively the natural number, δ is the sensitivity threshold of controller input, namely the minimum threshold of controller input, and the meaning of other symbol is the as the previous mentioned above.

The above control algorithm is to be fused the wisdom of human being, and it simulates the control thinking process of the human. Its outstanding advantage is that the designer is without more prior knowledge for controlled complex object, and it is strong in robustness of control system, fast in response speed, sensitive in error change, smooth in operating response and high in control precision.

5 Simulation Experiment and Its Analysis

The uncertainty of system has various types, the complexity grade is also not the same, the type and grade of uncertainty is difficult to make the measure, there is not a determined standard, and therefore, it is very difficult to make the simulation in practical control system. Generally, it takes a more representative mathematical model, through changing the order number and parameters of system model, and then it can judge the control quality of controller by means of validating the strong robustness of control algorithm. If it is the strong robust, then the control strategy or algorithm is advisable. Without loss of generality, it takes the two-order model with time lag as an example and the PID control algorithm as reference in the simulation, it adopt strong pulse signal interference to replace the response influence of model order, parameter and external disturbance, and then it examines the system response characteristic for control strategy or control algorithm. If it is much better than PID control algorithm in the respects of robustness performance, adjusting time, response rise time, overshoot performance and steady-state control accuracy, then the control strategy or algorithm is desirable. Assuming that the controlled object is a two-order system with time delay, the system parameter and time delay is changeable or uncertain, and the model is as the following.

$$G(s) = \frac{K_0 \times e^{-\tau s}}{(T_1 s + 1)(T_2 s + 1)}$$

For convenience of the simulation, it takes τ = 10 s, K_0 = 4.134, T_1 = 1.0 s, T_2 = 2.0 s in the model. In order to compare the control effect, it respectively adopts the PID and fusion HSIC control algorithm to make the control for a same control

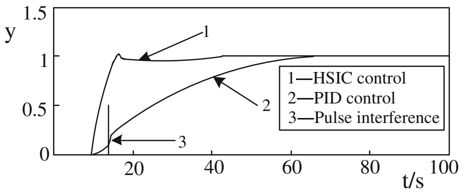


Fig. 4 Response under external disturbance

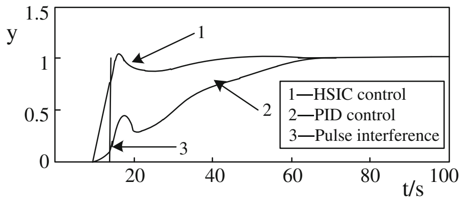


Fig. 5 Response under strong disturbance

object of two-order system with time lag. Under the condition of strong pulse signal interference, the pulse amplitude is respectively 0.5 and 1, and the pulse width is 0.2 s. The system joins the above strong pulse interference signal at $t = 15$ s, and the Fig. 4 and Fig. 5 is the contrast response curve respectively by PID and fusion HSIC control algorithm under different strong pulse interference signal, and compared with the response curve of PID control algorithm, the fusion control strategy (HSIC) shows very strong robustness. Under the strong pulse interference, the response of fusion HSIC almost does not change, and it is quick in rise time, short in transient time, relatively stable and smooth in response curve and high in steady-state control accuracy. According to the error size, error change direction and its change trend, the control algorithm of fusion HSIC makes adjust for controlled process at each control period, it forces the system error toward to zero, and therefore, it improves the response speed, shortens the transition time, suppresses the large amplitude overshoot of the system and avoids to induce the oscillation from over adjust. The simulation demonstrates its strong robustness of multi-modal based on human-simulated intelligence for uncertainty system.

6 Conclusions

The simulation results show that the HSIC fusion control strategy owns very good adaptability and strong robustness. The control strategy based on HSIC does not need to establish a strict mathematical model for controlled object, the demand is not high in prior knowledge for the system designer, and therefore, it is a more appropriate choice to uncertainty system control. The above experiment result under the strong pulse signal interference has demonstrated its adaptability, feasibility and effectiveness.

References

1. Li, T.F., Yang, Z., Sheng, C.Q., et al.: Analysis on correlated problem with control of uncertainty system. J. Chongqing Univ. (Nat. Sci. Ed.) **25**(7), 9–11 (2002)
2. Li, T.F., Feng, G.L., Zhong, B.X., et al.: Analysis on control strategy for a kind of uncertain complexity system. J. Chongqing Univ. (Nat. Sci. Ed.) **26**(1), 4–7 (2003)
3. Li, S.Y.: Fuzzy Control Neural Control and Intelligence Cybernetics. Harbin Institute of Technology Press, Harbin (2002)
4. Yi, J.K.: Intelligent Control Technology. Beijing University of Technology Press, Beijing (2004)
5. Li, Z.S., Tu, Y.Q.: Human Simulated Intelligent Controller. National Defense Industry Press, Beijing (2003)

Combined Interpolation Method for Single-Well Monitoring Data

Xiang-jun Xie, Meng-chuan Zhao and Liu-li Lu

Abstract Obtain accurate, reliable and complete single-well monitoring data is important task of oilfield dynamic monitoring. At present, oilfield dynamic monitoring technology has become more and more mature, but it is impossible to cover every single-wells or every horizons. Thus, the single-well monitoring data is not continuous, and the missing phenomenon is widespread. To solve this problem, this paper proposes a new combined interpolation algorithm based on SPSS which presents five missing data processing methods and cubic spline interpolation, and applied to a single-well monitored missing data processed. The numerical experiment shows that the new algorithm has the highest precision, and it as an effective approach for interpolating the missing data of single-well monitoring.

Keywords Single-well monitoring data · Missing data · SPSS · Cubic spline interpolation · Combined interpolation

1 Introduction

There are a lot of missing data phenomenon in real life. Bowley first proposed issue of missing data in 1915 [1]. After that, scholars have done a lot of problems with missing data research, and put forward many methods to fill the missing data. Missing data processing method is roughly divided into four categories: the first category is weighted method; second category is the case deletion method; third category is based on the interpolation method; The fourth category is model-based approach. On the interpolation method, can also be divided into single interpolation

X. Xie (✉) · M. Zhao
School of Sciences, Southwest Petroleum University, Chengdu 610500, China
e-mail: xiangjunxie@126.com

L. Lu
Information and Computing Sciences, Chengdu Technological University,
Chengdu 610500, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_36

method [2] and multiple interpolation method [3, 4]. Single interpolation method only construct an alternative value to each missing data, it based on the estimated. After the substitution of missing data, the corresponding statistical analysis was carried out on the new data. Rubin [5] first proposed multiple interpolation on the basis of a single interpolation in 1976. It was constructed to more than one alternative values for each missing value. Thus, a number of complete data sets will appear. Respectively, using the same method to give a number of processing results for each complete data set. Finally, the combination of these processing results can be obtained an estimated value of the target variable. Moreover, with the development of computer technology, there has been a lot of handling missing data statistical software such as R software, SPSS and SAS, which are simple to operate, easy to apply. Meanwhile, a large number of modern statistical methods applied to the problem of missing data also led to the development of this field and make handling missing data methods continue to improve.

Dynamic monitoring is an important basic work in the process of oilfield development, it always throughout the whole oilfield. Dynamic monitoring plays a crucial role, especially to improve the recovery of oilfield development goals. However, it is impossible for each oil well or each horizons are monitored due to the complexity of the monitoring work, it makes a lot of monitoring data are incomplete, missing data caused a lot of inconvenience to follow-up study work of the oilfield. Therefore, we need to find effective ways to fill in missing data to reflect the monitoring indicators information of oilfield, and lay a foundation for accurately predict oilfield development.

At present, using a single interpolation method for missing data of single-well, such as: mean interpolation, random interpolation, hot deck interpolation, cold deck interpolation and deductive interpolation, Lagrange interpolation and the cubic spline interpolation [6–9]. Single interpolation method can not be corrected inappropriate results and impact to subsequent research when appear inappropriate interpolation results. To overcome this deficiency of single interpolation algorithm, this paper proposes a new combined interpolation algorithm based on five kinds of missing data processing methods of SPSS software and cubic spline interpolation for single-well monitoring missing data. Numerical experiments show that the proposed method of interpolation of data with high accuracy and it as an efficient method for processing a missing data of single-well monitoring.

2 SPSS and Cubic Spline Interpolation Principle

2.1 SPSS Missing Data Processing Method

The SPSS statistical analysis software is recognized as three major data analysis software (SAS, SPSS and SYSTAT), it is powerful, user-friendly and widely used in various fields, the basic module provides five kinds of interpolation method about missing data.

(1) Series mean: with a mean of the entire column to impute missing data.
(2) Mean of nearby points: with missing values are close to the average number of points to impute missing values, specifically by the span of a few points near the point set.
(3) Median of nearby points: with missing values at several points near the median to impute missing values, specifically by the span of a few points near the point set.
(4) Linear interpolation: establish linear interpolation function by before and after data of missing value, and to calculate the interpolation of missing values by function approximation.
(5) Linear trend at point: set up linear regression equation, then calculated approximation of missing values using the equation.

2.2 Cubic Spline Interpolation

The cubic spline interpolation is a method for fitting a smooth curve of each main point, it is to form a smooth curve with connecting all the main points by constructing a polynomial. It has high convergence, high stability and interpolation function with good smoothness. And it has been widely used in engineering. The definition of cubic spline interpolation function is given below [10].

If the function $S(x) \in C^2[a, b]$, and it is a cubic polynomial in the interval $[x_j, x_{j+1}]$, where $a = x_0 < x_1 < \cdots < x_n = b$ are given nodes, claimed $S(x)$ is a cubic spline function on node $x_0, x_1, \ldots, x_n$. If given function value $y_j (j = 0, 1, \ldots, n)$ on the node x_j, and set up

$$S(x_j) = y_j \quad (j = 0, 1, \ldots, n) \tag{1}$$

We claimed $S(x)$ is a cubic spline function and obtained the formula cubic spline interpolation function by natural boundary conditions and three-corner method. Calculated as follows:

$$\begin{bmatrix} 2 & \lambda_0 & 0 & \cdots & 0 & 0 & 0 \\ \mu_1 & 2 & \lambda_1 & \cdots & 0 & 0 & 0 \\ 0 & \mu_2 & 2 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \mu_{n-1} & 2 & \lambda_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & \mu_n & 2 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix} \tag{2}$$

where, $h_j = x_j - x_{j-1}, \lambda_0 = \mu_0 = d_0 = d_n = 0,$ $\mu_j = \frac{h_{j-1}}{h_{j-1} + h_j}, \lambda_j = \frac{h_j}{h_{j-1} + h_j}, d_j = 6\frac{f[x_j, x_{j+1}] - f[x_{j-1}, x_j]}{h_{j+1} + h_j} = 6f[x_{j-1}, x_j, x_{j+1}]$, $f[x_j, x_{j+1}]$ is an order difference quotient, $f[x_{j-1}, x_j, x_{j+1}]$ is the second-order difference quotient. Linear equations can be solved by catching method [10], Solved $M_0, M_1, \ldots, M_n$ into formula (3), we can obtain cubic spline function.

$$\begin{aligned} S(x) = M_j \frac{(x_{j+1} - x)^3}{6h_j} + M_{j+1} \frac{(x - x_j)^3}{6h_j} + (y_j - \frac{M_j h_j^2}{6}) \frac{x_{j+1} - x}{h_j} \\ + (y_{j+1} - \frac{M_{j+1} h_j^2}{6}) \frac{x - x_j}{h_j} \quad (j = 0, 1, \ldots, n-1) \end{aligned} \tag{3}$$

3 Combined Interpolation Algorithm Model

Combined interpolation algorithm is an effective method of handling missing data, it can compensate for the lack of a single interpolation algorithm. In this paper, we propose a new combined interpolation algorithm based on SPSS and cubic spline interpolation, and use it to fill in the missing data of a single-well. The following principles are given a combination of interpolation algorithm.

First, the historical data is divided into two categories: the train and test sets. Secondly, we use five kinds of methods to provide missing data of SPSS and cubic spline interpolation based on the train set data and fill data of test set. Then, according to the mean absolute percentage error based on the interpolated value and the true value of test set, we can obtain the weights of each algorithm in combined interpolation algorithm. Finally, we get the final interpolated value by the six kinds of missing data interpolation algorithm and their weights.

Specific steps are as follows:
Let $x(t_i)(i = 1, 2, \ldots, n)$ represent monitor value of the single-well monitoring indicator x at time t_i. The lack of monitoring indicators data $x(t_j)$ at time t_j, $(j = 1, 2, \ldots, m)$.

Step 1 **Divide the train set and test set**

According to a certain proportion of the existing monitoring data for x randomly divided into a train set $x(t_p), (p = 1, 2, \ldots, k)$ and a test set $x(t_q)$, $(q = 1, 2, \ldots, h)$.

Step 2 **SPSS interpolation**

We impute missing data $x(t_j)$, $(j = 1, 2, \ldots, m)$ and test data $x(t_q)(q = 1, 2, \ldots, h)$ using five major functional modules of SPSS software and according to the train set data $x(t_p)(p = 1, 2, \ldots, k)$, interpolated values were recorded as:

$$x^1(t_j), x^2(t_j), x^3(t_j), x^4(t_j), x^5(t_j), (j = 1, 2, \ldots, m)$$
$$x^1(t_q), x^2(t_q), x^3(t_q), x^4(t_q), x^5(t_q), (q = 1, 2, \ldots, h)$$

Step 3 **Cubic spline interpolation**
We establish cubic spline interpolation function $S(x)$ according to the train set data $x(t_p)$ and use function values of the missing data and test set data to fill in missing data and test set data.
Interpolated values were recorded as follows:

$$x^6(t_j) = S(t_j)(j = 1, 2, \ldots, m), x^6(t_q) = S(t_q)(q = 1, 2, \ldots, h) \tag{4}$$

Step 4 **Error calculation**
Calculate the mean absolute percentage error (*MAPE*) between the real value and the interpolation value of test set according to formula (5), and then evaluate the accuracy of six kinds of interpolation methods:

$$MAPE_i = \frac{1}{h}\sum_{q=1}^{h}\left|\frac{x(t_q) - x^i(t_q)}{x(t_q)}\right|, i = 1, 2, \ldots, 6 \tag{5}$$

Step 5 **Determine the weights of combined interpolation algorithm**
Determining the weights of combined interpolation algorithm by *MAPE* for each method. The larger the *MAPE*, the smaller share of the weight. Weight is calculated as follows:

$$w_i = \frac{1/MAPE_i}{\sum_{i=1}^{6} 1/MAPE_i}, i = 1, 2, \ldots, 6 \tag{6}$$

Step 6 **Imputation of missing values**
According to formula (7) and using the weighted average value of interpolation value by six methods described as the final value of the missing data.

$$x(t_j) = \sum_{i=1}^{6} w_i x^i(t_j), j = 1, 2, \ldots, m \tag{7}$$

The proposed combination of interpolation algorithm to achieve a single-well monitoring data imputation of missing data. The proportion of the train and test sets can be set according to the actual situation, when the number of data is large, we can increase the proportion of the test set; When the number of data is small, we can reduce the proportion of the test set. The purpose of the test set is designed to calculate the *MAPE* of the algorithms to determine its weights in the combined interpolation algorithm, and realize the real meaning of the combination interpolation.

4 Application

We select 30 flow pressure values from January 2011 to September 2014 of Shengli Oilfield's YX4 well in this paper. Let 25 monitoring data as a train set, and the remaining five monitoring data as a test set, shown in Fig. 1.

Next, we impute missing the 15 monitoring data in Fig. 1. First, Using the five kinds of missing data processing methods provided by SPSS and cubic spline interpolation to fill the test set data and missing data, the data after interpolation shown in Fig. 2. Where, YX4_1 represents the interpolation value by series mean method; YX4_2 represents the interpolation value by mean of nearby points method; YX4_3 represents data by median of nearby points method; YX4_4 obtain by linear interpolation method; YX4_5 represents linear trend at point; YX4_6 represents data interpolation by way of cubic spline interpolation.

Table 1 shows the interpolated values of test set by six kinds of interpolation algorithm. The *MAPE* by the formula (5), we calculated the weights of six kinds of interpolation algorithms in the combined interpolation algorithm by the formula (6), the results shown in Table 2. Then, using Eq. (7) can be get the interpolated values of test set by combined interpolation algorithm and we can also get the *MAPE* by the formula (5). The results shown in Table 3.

Six interpolation method shown in Table 2, the *MAPE* of median of nearby points method (YX4_3) is the smallest value (0.07 %), it is the largest share of the weight (0.2551). The *MAPE* of series mean method (YX4_1) is the largest value (0.92 %), it is the smallest share of the weight (0.0182). In general, the smaller the *MAPE*, which is the larger share of the weight in combined interpolation algorithm.

Observed in Tables 2 and 3, we can see that the *MAPE* of combined interpolation algorithm is smaller than aforementioned six kinds of interpolation method's. Therefore, the combined interpolation algorithm is better than each single interpolation algorithm, it can be more accurate interpolation of missing data.

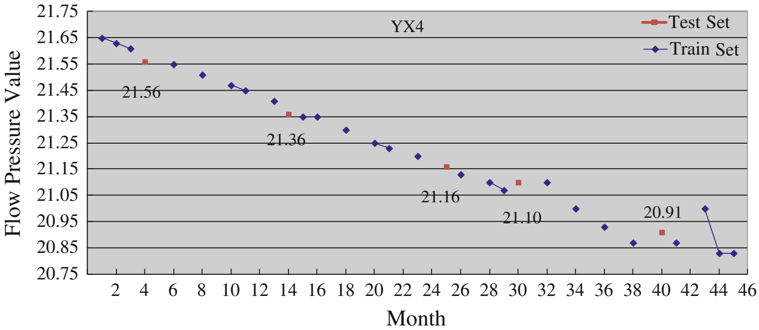


Fig. 1 Monitoring data of YX4 well's flow pressure values

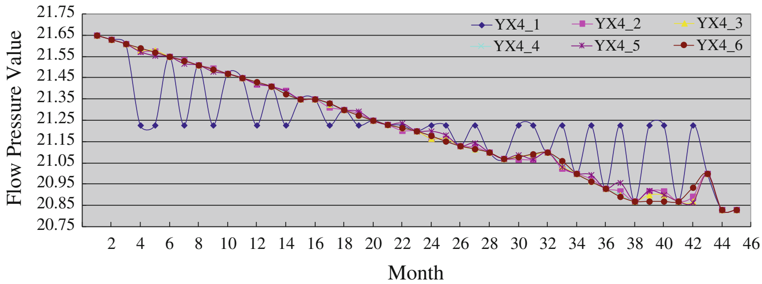


Fig. 2 Flow pressure values after interpolation of YX4 well by six methods

Table 1 Real value and six interpolation method's interpolated values of the test set

Real value	YX4_1	YX4_2	YX4_3	YX4_4	YX4_5	YX4_6
21.56	21.23	21.58	21.58	21.59	21.57	21.59
21.36	21.23	21.39	21.38	21.38	21.39	21.37
21.16	21.23	21.17	21.17	21.15	21.18	21.15
21.10	21.23	21.07	21.09	21.08	21.09	21.08
20.91	21.23	20.92	20.90	20.87	20.90	20.87

Table 2 MAPE and weights of six interpolation method

	YX4_1	YX4_2	YX4_3	YX4_4	YX4_5	YX4_6
MAPE (%)	0.92	0.08	0.07	0.11	0.08	0.11
w	0.0182	0.1981	0.2551	0.1525	0.2202	0.1559

Table 3 Combination interpolated values and MAPE

Real value	21.56	21.36	21.16	21.10	20.91	MAPE
Interpolated values	21.57	21.38	21.17	21.08	20.90	0.06 %

5 Conclusion

Oilfield monitoring data is incomplete due to various reasons, the researchers should use the appropriate method based on existing monitoring data to fill the missing data and it can bring convenience to follow-up research of oilfield. We present a combined interpolation algorithm to fill missing data in paper, which overcomes the shortcomings of single interpolation algorithm, and it makes full use of the advantages of different interpolation methods to enhance the interpolation results. Finally, numerical experiments show that the proposed method of interpolation of data with high accuracy. Additionally, SPSS software widely used and it has five kinds of missing data processing method with easy operation, simple and feasible. Cubic spline interpolation method is also easy to program implementation at the same time. In this paper, the combined interpolation algorithm has stronger practicability and as missing data processing method which can be widely applied to other areas.

References

1. Bowley, A.L.: An Elementary Manual of Statistics. PS King & Son, Limited (1915)
2. Tian, B.: Single imputation methods of missing data. Yingshan Acad. J. (Sci. Technol. Ed.) **25**(3), 17–19 (2011)
3. Pang, X.: Missing data multiple interpolation algorithm of processing method. Stat. Decis. (11), 88–90 (2012)
4. Pang, X.: A comparative study of missing data interpolation processing method. Stat. Decis. (24), 18–22 (2012)
5. Rubin, D.B.: Inference and missing data. Biometrika **63**(3), 581–592 (1976)
6. Velicer, W.F., Colby, S.M.: A comparison of missing-data procedures for ARIMA time-series analysis. Educ. Psychol. Measur. **65**(4), 596–615 (2005)
7. Muteki, K., MacGregor, J.F., Ueda, T.: Estimation of missing data using latent variable methods with auxiliary information. Chemometr. Intell. Lab. Syst. **78**(1), 41–50 (2005)
8. Hu, X., Chen, X., Qian Y., et al.: Research on the method of filling missing data in data processing. J. Hubei Univ. Technol. **28**(5), 82–84 (2013)
9. Zhao, Z., Liu, D., Jiang, H., et al.: Application of performance monitoring indices in predicting development indices of block. J. Southwest Pet. Univ. (Sci. Technol. Ed.) **31**(1), 116–120 (2009)
10. Li, Q., Wang, N., Yi, D.: Numerical Analysis. Beijing, Tsinghua University Press (2009)

Probability Distribution of T-System Based on B-P and L-S Implication

Yu-Bin Zhong , Yuan-zhi Deng , Zeng-liang Liu and Xue-hai Yuan

Abstract Fuzzy methods are widely used in the study of trustworthiness. Based on this fact, the paper researches the trustworthiness system (T-System) and probability distribution theory based on bounded product (B-P) implication and Larsen square (L-S) implication. Firstly, we convert a group of single-input and single-output data into fuzzy inference rules and generate fuzzy relation by selecting the appropriate fuzzy implication operator, then calculate joint probability density function of two-dimensional random variables by using of this fuzzy relation. Two specific probability density functions can be obtained by selecting the fuzzy implication as bounded product implication or Larsen square implication. Secondly, we study of the center-of-gravity trustworthiness system based on these two probability distributions. Finally, we gave the sufficient conditions of universal approximations for those trustworthiness systems.

Keywords T-system · Probability distribution · Universal approximation · Application research

---

Y.-B. Zhong (✉) · Y. Deng
Key Laboratory of Mathematics and Interdisciplinary Sciences of Guagdong Higher Education Institutes, College of Mathematics and Information Sciences, Guangzhou University Guangzhou, Guangdong 510006, People's Republic of China
e-mail: Zhong_yb@163.com

Y. Deng
e-mail: 34930995@qq.com

Y.-B. Zhong
College of Mathematics and Information Sciences, Guangzhou University Guangzhou, Guangzhou, China

Z. Liu
Institute of Information Operation National Defence University of PLA, Beijing 100091, People's Republic of China
e-mail: chinapd@tom.com

X. Yuan
Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China
e-mail: yuanxuehai@yahoo.com.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_37

1 Introduction

Fuzzy methods have wide application backgrounds in the study of trustworthiness, for example, (1) a software is unique, and trustworthiness behaviors of software are fuzzy in nature; (2) in control systems, we may have fuzzy control, design trustworthiness and intelligent control, etc.; (3) in expert system research, it is necessary to take into consideration of the consistency, completeness, independence and redundancy of knowledge bases, but it is also important to consider fuzzy expert systems and the spread of non-trustworthiness factors. Therefore, it is significant to study the trustworthiness system and its probability distribution theory, and this will benefit a lot to production and life practices.

Construction of trustworthiness systems and the universal approximation of them are hot topics in the research of trustworthiness system and trustworthiness control theory. In the construction of trustworthiness systems, the single fuzzifier of the input variable has been mentioned in most literatures until now. And fuzzy inference is CRI reasoning method [1, 2] or triple I method [3]. We must determine the fuzzy implication operator before constructing fuzzy inference relationship and fuzzy reasoning. The choice of fuzzy implication operator will have a huge impact on the trustworthiness system. References [4–7] point out that, CRI reasoning method and triple I method have impact on the trustworthiness system when only using the conjunction type fuzzy implication operators, such as Mamdani implication and Larsen implication. Because of involved some complex integral, researchers have not yet given the specific expression of the trustworthiness system when constructing them by using center-of-gravity method until now. How to obtain the specific expression of this trustworthiness system is one motivation in our paper.

As we know, the trustworthiness system constructed by center-of-gravity method can be approximately reduced to some forms of interpolation [8] and this trustworthiness system has meaning of probability theory. It is the best approximation of system under the sense of least-squares [9]. But how to determine the corresponding probability density function of fuzzy implication operator has not been resolved, and this is the another motivation in our paper.

This article is organized as follows: Sect. 2 is preliminary; Sect. 3 is the probability distributions and marginal probability distribution of bounded product implication and Larsen square implication; Sect. 4 gives the center-of-gravity method trustworthiness systems of these two probability distributions and the sufficient condition with universal approximation for these systems; Sect. 5 is a overlook of application analysis of trustworthiness systems and we have our conclusion in Sect. 6.

2 Preliminary

Suppose $\{(x_i, y_i)\}_{(1\le i\le n)}$ is a group of input-output data, then

$$a = x_1 < x_2 < \cdots < x_n = b, c = y_1 < y_2 < \cdots < y_n = d,$$

or

$$a = x_1 < x_2 < \cdots < x_n = b, c = y_n < y_{n-1} < \cdots < y_1 = d.$$

we construct a two-phase triangular wave by using these data, namely

$$A_1(x) = \begin{cases} \frac{x_2-x}{x_2-x_1}, \ x \in [x_1, x_2] \\ 0, other \end{cases}, \ A_n(x) = \begin{cases} \frac{x_n-x}{x_n-x_{n-1}}, \ x \in [x_1, x_2] \\ 0, other \end{cases},$$

for $i = 2, 3, \ldots, n-1$,

$$A_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}}, \ x \in [x_{i-1}, x_i] \\ \frac{x_{i+1}-x}{x_{i+1}-x_i}, \ x \in [x_i, x_{i+1}] \\ 0, other \end{cases}, \ B_i(y) = \begin{cases} \frac{y-y_{i-1}}{y_i-y_{i-1}}, \ y \in [y_{i-1}, y_i] \\ \frac{y_{i+1}-y}{y_{i+1}-y_i}, \ y \in [y_i, y_{i+1}] \\ 0, other \end{cases},$$

$$B_1(y) = \begin{cases} \frac{y_2-y}{y_2-y_1}, \ y \in [y_1, y_2] \\ 0, other \end{cases}, \ B_n(y) = \begin{cases} \frac{y_n-y}{y_n-y_{n-1}}, \ y \in [y_{n-1}, y_n] \\ 0, other \end{cases},$$

then x_i, y_i is the peak point of A_i and y_i, namely $A_i(x_i) = 1, B_i(y_i) = 1, (i = 1, 2, \ldots, n)$, and when $x \in [x_i, x_{i+1}], A_i(x) + A_{i+1}(x) = 1, A_j(x) = 0 (j \neq i, i+1)$, when $y \in [y_i, y_{i+1}]$, $B_i(y) + B_{i+1}(y) = 1, B_j(y) = 0 (j \neq i, i+1)$.
Then we have fuzzy inference rule

$$if\ x\ is\ A_i,\ then\ y\ is\ B_i (i = 1, 2, \ldots, n). \tag{1}$$

Let θ be the fuzzy implication operator, then from (1) can get the fuzz relationship $R(x, y) = \bigvee_{i=1}^{n} \theta(A_i(x), B_i(y))$.
Let $q(x, y) = \begin{cases} R(x, y), (x, y) \in X \times Y \\ 0, other \end{cases}$, and $H(2, n, \theta, \vee) \triangleq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} q(x, y) dxdy$, when $X = [a, b], Y = [c, d]$, we have

$$H(2, n, \theta, \vee) = \int_a^b \int_c^d R(x, y) dxdy. \tag{2}$$

we call $H(2, n, \theta, \vee)$ H function which has parameter $2, n, \theta, \vee$, and “2” indicates $q(x, y)$ is binary function; n means the amount of inference rules; θ is fuzzy implication operator, and $\vee$ = “*max*”.

If $H(2, n, \theta, \vee) > 0$, let

$$f(x,y) \triangleq \frac{q(x,y)}{H(2,n,\theta,\vee)}, \tag{3}$$

clearly, (a) $f(x,y) \geq 0$; (b) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$. So, $f(x,y)$ can be regarded as a random vector of the joint probability density function (ξ, η) [9].

Suppose $R(x,y)$ is fuzzy relationship determined by inference rules (1), and $A^*(x)$ is a fuzzy single point of input variables, namely $A^*(x^{'}) = \begin{cases} 1, x^{'} = x \\ 0, x^{'} \neq x \end{cases}$. Let $B^* = A \circ R$, namely $B^*(y) = \underset{x^{'} \in X}{\vee} (A^*(x^{'}) \wedge R(x^{'}, y)) = R(x,y)$. Then

$$\bar{S}(x) = \frac{\int_c^d yB^*(y)dy}{\int_c^d B^*(y)dy} \tag{4}$$

is a center-of-gravity method trustworthiness system [8].

3 The Probability Distribution of Several Single-Input and Single-Output Trustworthiness System

If $\{(x_i, y_i)\}_{1 \leq i \leq n}$ satisfies:

$$a = x_1 < x_2 < \cdots < x_n = b, c = y_1 < y_2 < \cdots < y_n = d.$$

Establishing fuzzy relations according to the fuzzy inference rule (1) is one of the four processes of constructing a trustworthiness system. If using the traditional method, then the fuzzy relation is $R(x,y) = \overset{n}{\underset{i=1}{\vee}} \theta(A_i(x), B_i(y))$. Since the fuzzy implication is conjunction implication, we have the following result:

if $x \in [x_i, x_{i+1}]$, then $R(x,y) = \theta(A_i(x), B_i(y)) \vee \theta(A_{i+1}(x), B_{i+1}(y))$.

Because the data is monotonic, we make some adjustments to the fuzzy reasoning relationship: when $x \in [x_i, x_{i+1}]$, let

$$R(x,y) = \begin{cases} \theta(A_i(x), B_i(y)) \vee \theta(A_{i+1}(x), B_{i+1}(y)), & y \in [y_i, y_{i+1}] \\ 0, other. \end{cases} \tag{5}$$

We discuss the reasoning relationship determine by (5), then we have the following probability results:

Theorem 1 *If* $\theta(a,b) = T_m(a,b) = (a+b-1) \vee 0$, *let* $H = \frac{1}{3}\sum_{i=1}^{n-1}(x_{i+1} - x_i)(y_{i+1} - y_i)$. $D_i = \{(x,y) \in I_i \times J_i | y \leq A_{i+1}(x)y_i + A_i(x)y_{i+1}\}, E_i = I_i \times J_i - D_i$, *then*

(1) Probability density function $f(x, y)$ (called Bounded plot distribution) determined by formula (3) is

$$f(x,y)=\begin{cases}\frac{1}{H}(A_i(x)+B_i(y)-1),\ (x,y)\in D_i\\ \frac{1}{H}(A_{i+1}(x)+B_{i+1}(y)-1),\ (x,y)\in E_i\\ 0, other\end{cases} \tag{6}$$

(2) Marginal distribution functions of $f(x, y)$ are

$$f_\xi(x)=\frac{1}{2T}(y_{i+1}-y_i)(1-2A_i(x)A_{i+1}(x)),\ x\in[x_i,x_{i+1}](i=1,2,\dots n). \tag{7}$$

$$f_\eta(x)=\frac{1}{2T}(x_{i+1}-x_i)(1-2B_i(y)B_{i+1}(y)),\ y\in[y_i,y_{i+1}](i=1,2,\dots n). \tag{8}$$

Theorem 2 *If $\theta(a, b) = a^2 b$, then*

(1) Probability density function $f(x, y)$ (called Larsen square distribution) determined by formula (3) is

$$f(x,y)=\begin{cases}\frac{1}{C}A_i^2(x)B_i(y), \exists i,\ (x,y)\in F_i\\ \frac{1}{C}A_{i+1}^2(x)B_{i+1}(y), \exists i,\ (x,y)\in G_i\\ 0, other\end{cases} \tag{9}$$

(2) Marginal density function of $f(x, y)$ is: for $x \in [x_i, x_{i+1}](i = 1, 2, \dots, n)$

$$f_\xi(x)=\frac{1}{2C}(y_{i+1}-y_i)\frac{(1-A_i(x)A_{i+1}(x))(1-3A_i(x)A_{i+1}(x))}{1-A_i(x)A_{i+1}(x)}, \tag{10}$$

$$f_\eta(y)=\frac{1}{3C}(x_{i+1}-x_i)(1-\frac{B_i(y)B_{i+1}(y)}{1+2\sqrt{B_i(y)B_{i+1}(y)}})(y\in[y_i,y_{i+1}]), \tag{11}$$

where $F_i=\{(x,y)\in I_i\times J_i|y\le\frac{A_{i+1}^2(x)}{A_i^2(x)+A_{i+1}^2(x)}y_i+\frac{A_i^2(x)}{A_i^2(x)+A_{i+1}^2(x)}y_{i+1}\}, G_i=(I_i\times J_i)-F_i, K=\frac{1}{2}-\frac{1}{16}\pi,$ *and* $C=K\sum_{i=1}^{n-1}(x_{i+1}-x_i)(y_{i+1}-y_i).$

Note 1: In fact, Theorems 1 and 2 establish a method to construct probability distribution from the known data structures.

4 Center-of-Gravity Trustworthiness System

Let $f(x, y)$ be the probability density function obtained in our study, we can know that from the discussion above,

$$\overline{S}(x) = \frac{\int_c^d yB^*(y)dy}{\int_c^d B^*(y)dy} = \frac{\int_c^d yR(x,y)dy}{\int_c^d R(x,y)dy} = \frac{\int_{-\infty}^{+\infty} yf(x,y)dy}{\int_{-\infty}^{+\infty} f(x,y)dy}, \tag{12}$$

so we have the following theorem.

Theorem 3 *If* $\theta =$ *Bounded Product implication, where* $\overline{S}_T(x) = A_i^*(x)y_i + A_{i+1}^*(x)y_{i+1}$, *where*

$$A_i^*(x) = \frac{\frac{2}{3}A_{i+1}^3(x) - A_{i+1}(x) + \frac{2}{3}}{1 - 2A_i(x)A_{i+1}(x)} \tag{13}$$

$$A_{i+1}^*(x) = 1 - A_i^*(x) = \frac{\frac{2}{3}A_i^3(x) - A_i(x) + \frac{2}{3}}{1 - 2A_i(x)A_{i+1}(x)} \tag{14}$$

and $\overline{S}_T(x_i) = \frac{2}{3}y_i + \frac{1}{3}y_{i+1}, \overline{S}_L(x_{i+1}) = \frac{1}{3}y_i + \frac{2}{3}y_{i+1}$.

Proof From the theorem above we can know that:

$$\int_{-\infty}^{+\infty} f(x,y)dy = f_\xi(x) = \frac{y_{i+1} - y_i}{T}[\frac{1}{2} - A_i(x)A_{i+1}(x)]$$

and

$$\int_{-\infty}^{+\infty} yf(x,y)dy = \frac{1}{T}[\int_{y_i}^{y_i^*} y(A_i(x) + B_i(x) - 1)dy + \int_{y_i^*}^{y_{i+1}} y(A_{i+1}(x) + B_{i+1}(x) - 1)dy]$$
$$= \frac{y_{i+1}-y_i}{T}[(A_i^2(x) - \frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{6})y_i + (\frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{3})y_{i+1}]$$

so

$$\overline{S}_T(x) = \frac{\int_{-\infty}^{+\infty} yf(x,y)dy}{\int_{-\infty}^{+\infty} f(x,y)dy} = \frac{(A_i^2(x) - \frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{6})y_i}{\frac{1}{2} - A_i(x)A_{i+1}(x)} + \frac{(\frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{3})y_{i+1}}{\frac{1}{2} - A_i(x)A_{i+1}(x)}$$
$$= A_i^*(x)y_i + A_{i+1}^*(x)y_{i+1},$$

and

$$A_i^*(x) + A_{i+1}^*(x) = \frac{1}{\frac{1}{2} - A_i(x)A_{i+1}(x)}[A_i^2(x) - \frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{6} + \frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x)$$
$$+ \frac{1}{3}] = \frac{\frac{1}{2} - A_i(x)(1 - A_i(x))}{\frac{1}{2} - A_i(x)A_{i+1}(x)} = 1.$$

And because
$A_i^2(x) - \frac{1}{3}A_i^3(x) - \frac{1}{2}A_i(x) + \frac{1}{6} = (1 - A_{i+1}(x))^2 - \frac{1}{3}(1 - A_{i+1}(x))^3 - \frac{1}{2}(1 - A_{i+1}(x)) + \frac{1}{6} = \frac{1}{3}A_{i+1}^3(x) - \frac{1}{2}A_{i+1}(x) + \frac{1}{3}$
we have

$$A_i^*(x) = \frac{\frac{2}{3}A_{i+1}^3(x) - A_{i+1}(x) + \frac{2}{3}}{1 - 2A_i(x)A_{i+1}(x)}, A_{i+1}^*(x) = \frac{\frac{2}{3}A_i^3(x) - A_i(x) + \frac{2}{3}}{1 - 2A_i(x)A_{i+1}(x)}.$$

We study the trustworthiness system exported by Larsen square distribution below.

Theorem 4 $\overline{S}_{L^2}(x) = \frac{\int_{-\infty}^{+\infty} yf(x,y)dy}{\int_{-\infty}^{+\infty} f(x,y)dy} = C_i^*(x)y_i + C_{i+1}^*(x)y_{i+1}$, *where*

$$C_{i+1}^*(x) = \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x) + A_{i+1}^2(x))^2}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^2(x))(A_i^2(x) + A_{i+1}^2(x))}, \tag{15}$$

$C_i^*(x)y_i + C_{i+1}^*(x)y_{i+1} \equiv 1$ *and* $\overline{S}_{L^2}(x_i) = \frac{2}{3}y_i + \frac{1}{3}y_{i+1}, \overline{S}_{L^2}(x_{i+1}) = \frac{1}{3}y_i + \frac{2}{3}y_{i+1}$.

Proof

$$\begin{aligned}
\int_{-\infty}^{+\infty} f(x,y)dy &= \int_{y_i}^{y_{i+1}} f(x,y)dy = \frac{y_{i+1} - y_i}{2C} \frac{(1 - A_i(x)A_{i+1}(x))(1 - 3A_i(x)A_{i+1}(x))}{1 - 2A_i(x)A_{i+1}(x)}, \\
\int_{-\infty}^{+\infty} yf(x,y)dy &= \frac{1}{C}[\int_{y_i}^{\hat{y}_i} yA_i^2(x)B_i(y)dy + \int_{\hat{y}_i}^{y_{i+1}} yA_{i+1}^2(x)B_{i+1}(y)dy] \\
&= \frac{y_{i+1} - y_i}{C}[[A_i^2(x)(A_i^*(x) - (A_i^*(x))^2 + \frac{1}{3}(A_i^*(x))^3) \\
&\quad + A_{i+1}^2(x)(\frac{1}{2}(1 - A_i^*(x))^2) - \frac{1}{3}(1 - (A_i^*(x))^3))]y_i \\
&\quad + [A_i^2(x)(\frac{1}{2}(A_i^*(x))^2 - \frac{1}{3}(A_i^*(x))^3) + \frac{1}{3}A_{i+1}^2(x)(1 - (A_i^*(x))^3)]y_{i+1}] \\
&= \frac{y_{i+1} - y_i}{C}[D_i(x)y_i + D_{i+1}(x)y_{i+1}].
\end{aligned}$$

Then

$$\begin{aligned}
D_i(x) + D_{i+1}(x) &= A_i^2(x)[A_i^*(x) - \frac{1}{2}(A_i^*(x))^2] + \frac{1}{2}A_{i+1}^2(x)[1 - (A_i^*(x))^2] \\
&= A_i^2(x)A_i^*(x) + \frac{1}{2}A_{i+1}^2(x) - \frac{1}{2}(A_i^*(x))^2(A_i^2(x) + A_{i+1}^2(x)) \\
&= \frac{2A_i^4(x) + A_{i+1}^2(x)(A_i^2(x) + A_{i+1}^2(x)) - A_i^4(x)}{2(A_i^2(x) + A_{i+1}^2(x))} \\
&= \frac{(1 - A_i(x)A_{i+1}(x))(1 - 3A_i(x)A_{i+1}(x))}{2(1 - 2A_i(x)A_{i+1}(x))}.
\end{aligned}$$

So, $C_i^*(x) + C_{i+1}^*(x) = 1$. And

$$
\begin{aligned}
D_{i+1}(x) &= A_i^2(x)[\frac{1}{2}(A_i^*(x))^2 - \frac{1}{3}(A_i^*(x))^3] + \frac{1}{3}A_{i+1}^2(x)(1-(A_i^*(x))^3) \\
&= \frac{3A_i^2(x)A_i^4(x)(A_i^2(x)+A_{i+1}^2(x)) - 2A_i^8(x) + 2A_{i+1}^2(x)((A_i^2(x)+A_{i+1}^2(x))^3 - A_i^6(x))}{6(A_i^2(x)+A_{i+1}^2(x))^3} \\
&= \frac{A_i^8(x) + 3A_i^6(x)A_{i+1}^2(x) + 2A_{i+1}^8(x) + 6A_i^4(x)A_{i+1}^4(x) + 6A_i^2(x)A_{i+1}^2(x)}{6(A_i^2(x)+A_{i+1}^2(x))^3} \\
&= \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x)+A_{i+1}^2(x))^2}{6(A_i^2(x)+A_{i+1}^2(x))^2},
\end{aligned}
$$

thus

$$
\begin{aligned}
C_{i+1}^*(x) &= \frac{2D_{i+1}(x)}{\dfrac{A_i^4(x)+A_i^2(x)A_{i+1}^2(x)+A_{i+1}^4(x)}{(A_{i+1}^2(x)+A_i^2(x))}} \\
&= \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x)+A_{i+1}^2(x))^2}{3(A_i^4(x)+A_i^2(x)A_{i+1}^2(x)+A_{i+1}^4(x))(A_i^2(x)+A_{i+1}^2(x))} \\
&= \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x)+A_{i+1}^2(x))^2}{3(1-A_i(x)A_{i+1}(x))(1-2A_i(x)A_{i+1}(x))(1-3A_i(x)A_{i+1}(x))}.
\end{aligned}
$$

So $\overline{S}(x) = C_i^*(x)y_i + C_{i+1}^*(x)y_{i+1}$,
$C_{i+1}^*(x_{i+1}) = \frac{2}{3}; C_i^*(x_{i+1}) = 1 - \frac{2}{3} = \frac{1}{3}; C_{i+1}^*(x_i) = \frac{1}{3}; C_i^*(x_{i+1}) = 1 - \frac{1}{3} = \frac{2}{3}$,
so $\overline{S}(x_i) = \frac{2}{3}y_i + \frac{1}{3}y_{i+1}, \overline{S}(x_{i+1}) = \frac{1}{3}y_i + \frac{2}{3}y_{i+1}$.

We study the universal approximations of trustworthiness systems $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$. Suppose $s(x)$ is a known system and $s(x_i) = y_i$. Let $h = \max\limits_{1\le i\le n-1} \Delta x_i$, $\|s\|_\infty = \max\limits_{x\in[a,b]} |s(x)|$. Suppose $F_1(x) = A_i(x)y_i + A_{i+1}(x)y_{i+1}, (x \in [x_i, x_{i+1}])$. From Ref. [10], we can know that: $\|s - F_1\| \le \frac{1}{8}\|s''\|_\infty h^2$. So we have

Theorem 5 *When* $\overline{S}(x) \in \{\overline{S}_T(x), \overline{S}_{L^2}(x)\}$,

$$
\|\overline{S}(x) - F_1(x)\| \le \frac{1}{3}\|s'\|_\infty h. \tag{16}
$$

Proof (1) When $\overline{S}(x) = \overline{S}_T(x)$,

$$
A_i^*(x) = \frac{\frac{1}{3}A_{i+1}^3(x) - A_{i+1}(x) + \frac{2}{3}}{1 - 2A_i(x)A_{i+1}(x)}.
$$

Then

$$A_i^*(x) - A_i(x) = \frac{\frac{1}{3}A_{i+1}^3(x) - A_{i+1}(x) + \frac{2}{3} - A_i(x) + 2A_i^2(x)A_{i+1}(x)}{1 - 2A_i(x)A_{i+1}(x)}$$
$$= \frac{2A_{i+1}^2(x) - 1 + 6A_i^2(x)A_{i+1}(x)}{3[1 - 2A_i(x)A_{i+1}(x)]}$$
$$= \frac{[1 - 2A_i(x)]^3}{3[1 - 2A_i(x)A_{i+1}(x)]}.$$

Then

$$A_{i+1}^*(x) - A_{i+1}(x) = -\frac{[1 - 2A_i(x)]^3}{3[1 - 2A_i(x)A_{i+1}(x)]},$$

$$\overline{S}_T(x) - F_1(x) = \frac{[1 - 2A_i(x)]^3}{3[1 - 2A_i(x)A_{i+1}(x)]}(y_i - y_{i+1}).$$

Then

$$|\overline{S}_T(x) - F_1(x)| = \frac{|[1 - 2A_i(x)]^3|}{3[1 - 2A_i(x)A_{i+1}(x)]}|y_{i+1} - y_i| \le \frac{1}{3}|y_{i+1} - y_i| \le \frac{1}{3}\|s'\|_\infty h.$$

(2)

$$C_{i+1}^*(x) = \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x) + A_{i+1}^2(x))^2}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x))(A_i^2(x) + A_{i+1}^2(x))},$$
$$A_{i+1}^*(x) = \frac{A_{i+1}^2(x)}{A_i^2(x) + A_{i+1}^2(x)}$$

then

$$C_{i+1}^*(x) - A_{i+1}^*(x) = \frac{A_i^6(x) + 2A_{i+1}^2(x)(A_i^2(x) + A_{i+1}^2(x))^2}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x))(A_i^2(x) + A_{i+1}^2(x))} -$$
$$\frac{A_{i+1}^2(x)}{A_i^2(x) + A_{i+1}^2(x)} = \frac{(A_i^2(x) - A_{i+1}^2(x)(A_i^4(x) + A_{i+1}^4(x))}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x))(A_i^2(x) + A_{i+1}^2(x))}.$$

$C_i^*(x) - A_i^*(x) = -(C_{i+1}^*(x) - A_{i+1}^*(x))$.
Then $\overline{S}_{L^2}(x) - F_1(x) = (C_{i+1}^*(x) - A_{i+1}^*(x))(y_{i+1} - y_i)$. So

$$|\overline{S}_{L^2}(x) - F_1(x)| = |C_{i+1}^*(x) - A_{i+1}^*(x)||y_{i+1} - y_i|$$
$$= \frac{|A_i^2(x) - A_{i+1}^2(x)|(A_i^4(x) + A_{i+1}^4(x))}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x))(A_i^2(x) + A_{i+1}^2(x))}|y_{i+1} - y_i| \le \frac{1}{3}|y_{i+1} - y_i| \le \frac{1}{3}\|s'\|_\infty h.$$

And

$$\begin{aligned} C^*_{i+1}(x) - A_{i+1}(x) &= C^*_{i+1}(x) - A^*_{i+1}(x) + A^*_{i+1}(x) - A_{i+1}(x) \\ &= \frac{(A_i^2(x) - A_{i+1}^2(x))(A_i^4(x) + A_{i+1}^4(x))}{3(A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x))(A_i^2(x) + A_{i+1}^2(x))} + \frac{A_i(x)A_{i+1}(x)[A_{i+1}(x) - A_i(x)]}{A_i^2(x) + A_{i+1}^2(x)} \\ &= \frac{[A_i^2(x) - A_{i+1}^2(x)]}{3[A_i^2(x) + A_{i+1}^2(x)]}[\frac{A_i^4(x) + A_{i+1}^4(x)}{A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x)} - 3A_i(x)A_{i+1}(x)] \\ &= \frac{[A_i^2(x) - A_{i+1}^2(x)]}{3[A_i^2(x) + A_{i+1}^2(x)]}[1 - \frac{A_i^2(x)A_{i+1}^2(x)}{A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x)} - 3A_i(x)A_{i+1}(x)] \end{aligned}$$

Because of $|\frac{[A_i^2(x) - A_{i+1}^2(x)]}{3[A_i^2(x) + A_{i+1}^2(x)]}| \leq 1$,
$1 \geq 1 - \frac{A_i^2(x)A_{i+1}^2(x)}{A_i^4(x) + A_i^2(x)A_{i+1}^2(x) + A_{i+1}^4(x)} - 3A_i(x)A_{i+1}(x) \geq 1 - \frac{1}{3} - \frac{3}{4} = \frac{5}{12} > 0.$
Thus $|C^*_{i+1}(x) - A_{i+1}(x)| \leq \frac{1}{3}$. So

$$|\overline{S}_{L^2}(x) - F_1(x)| = |C^*_{i+1}(x) - A_{i+1}(x)||y_{i+1} - y_i| \leq \frac{1}{3}\|s'\|_\infty h$$

Then $\|\overline{S}_{L^2}(x) - F_1(x)\|_\infty \leq \frac{1}{3}\|s'\|_\infty h.$

Theorem 6 *When* $\overline{S}(x) \in \{\overline{S}_T(x), \overline{S}_{L^2}(x)\}$,

$$\|s - \overline{S}\|_\infty \leq \frac{1}{8}\|s''\|_\infty h^2 + \frac{1}{3}\|s'\|_\infty h. \tag{17}$$

Proof $\|s - \overline{S}\|_\infty \leq \|s - F_1\|_\infty + \|\overline{S} - F_1\|_\infty \leq \frac{1}{8}\|s''\|_\infty h^2 + \frac{1}{3}\|s'\|_\infty h.$

Note 2 (1) We get $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$ are assumed $s(x)$ to be monotonic function. For non-monotonic function, we can divide X into several intervals, and let $s(x)$ are monotonic in each interval. Because $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$ can approximate $s(x)$ in each monotonic interval, $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$ can approximate $s(x)$ in entire X. For example, let $s(x) = \sin x, X = [-\pi, \pi], X_1 = [-\pi, -\frac{\pi}{2}], X_2 = (-\frac{\pi}{2}, -\frac{\pi}{2}), X_3 = [-\frac{\pi}{2}, -\pi]$. Then $s(x)$ is monotonic function in X_1, X_2, X_3. If $\varepsilon = 0.1$, then from Theorem 6 we can know that: $n = 17$ in X_1 and X_3, $n = 24$ in X_2. So we let $n = 17 \times 2 + 33 = 67$, then $\overline{S}_{L^2}(x)$ can approximate $s(x)$ with error not more than 0.1.
(2) From Theorem 6 we can know that: $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$ have first-order approximation accuracy to s and they have the same error estimates upper bound.

Example 1 Let $s(x) = \sin x, x \in [-3, 3]$. Then $\|S''\|_\infty = \|S'\|_\infty = 1$.
If $\varepsilon = 0.1$, from $\frac{1}{8}\|S''\|h^2 + \frac{1}{3}\|S'\|_\infty h < 0.1$ we can know that $n = 24$.

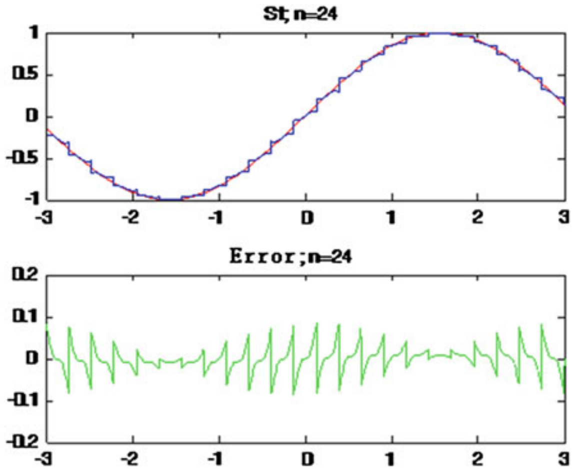


Fig. 1 The simulation result of $\overline{S}_T(x)$

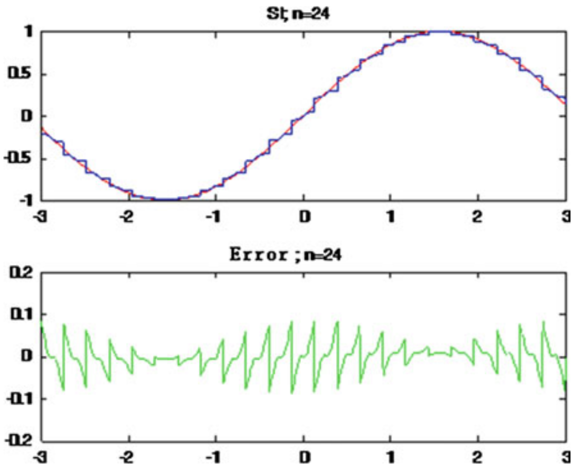


Fig. 2 The simulation result of $\overline{S}_{L^2}(x)$

For $\varepsilon = 0.1$, Figs. 1 and 2 show the simulation figure and the error estimates curves of $\overline{S}_T(x)$ and $\overline{S}_{L^2}(x)$ to s.

5 Application Analysis

The trustworthiness theory and systems in this paper have wide application fields. It provides a completely new thought in theory and application study on such fields as fuzzy logic and neural networks, factor neural networks, and fuzzy expert systems. Trustworthiness system, together with its theories, can be used in quantifiable description and measurement of trustworthy software [10–14], and this would be a new direction of expand research on software trustworthiness metrics models. From the view of trustworthiness theory and fuzzy system, network attack and defence, engineering system automatic design process and the design of fuzzy controller, etc., can also be explored, providing a new research method and direction [15] for these fields.

6 Conclusion

This paper has researched bounded product implication and Larsen square implication and obtained two specific probability density functions. And we got the center-of-gravity trustworthiness systems of these two probability distributions, then we gave sufficient condition of the universal approximations for those trustworthiness systems. As for application researches, trustworthiness system can be applied in software trustworthiness metrics, network attack and defence, engineering system automatic design process and the design of fuzzy controller, etc.

Acknowledgments This research was supported by Information and Computing Science Outstanding Talent Training project of Guangdong province (No.20153324).

References

1. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning I. Inf. Sci. **8**(3), 199–251 (1975)
2. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning II. Inf. Sci. **8**(4), 301–359 (1975)
3. Wang, G.J.: The triple I method with total inference rules of fuzzy reasoning.Science in China (Series E) **29**(1), 43–53 (1999)
4. Li, H.-X., You, F., Peng, J.: Fuzzy controllers based on some fuzzy implication operators and their response functions. Prog. Nat. Sci. **14**(1), 15–20 (2004)
5. Li, H.-X., Peng, J.Y., Wang, J.Y.: Fuzzy systems and their response functions based on commonly used fuzzy implication operators. Control Theory Appl. **22**(3), 341–347 (2005)
6. Li, H.-X., Peng, J.Y., Wang, J.Y.: Fuzzy systems based on triple I algorithm and their response ability. J. Syst. Sci. Math. Sci. **25**(5), 578–590 (2005)
7. Hou, J., You, F., Li, H.X.: Some fuzzy controllers based on the triple I algorithm and their response ability. Prog. Nat. Sci. **15**(1), 29–37 (2005)
8. Li, H.-X.: Interpolation mechanism of fuzzy control. Sci. China (Ser. E) **28**(3), 312–320 (1998)
9. Li, H.-X.: Probability representation of fuzzy systems. Sci. China (Ser. E) **36**(4), 373–397 (2006)
10. Tang, Y.X., Liu, Z.L.: Progress in software trustworthiness metrics models. Comput. Eng. Appl. **46**(27), 12–16 (2010)
11. Wang, Y., Liu, Z.: Risk assessment model for network security based on PRA. Comput. Eng. **32**(1), 40–42 (2006)
12. Liu, X., Liu, Z., Yu, D.: Research on method of security management metrics based on hierarchy protection. Control Manage. **45**, 39–40 (2007)
13. Liu, X., Liu, Z., Yu, D.: Method of security management metrics based on AHP model. Microcomput. Inf. **160** 33–34 (2007)

14. Wang, X., Chen, Y., Liu, Z.: Component-based software developing method for group companies management information system. Wuhan Univ. J. Nat. Sci. **13**(1), 37–44 (2008)
15. Wang, L.X.: A Course in Fuzzy Systems and Fuzzy Control. Prentice-Hall, Inc. (1997)

A New Design Technology for Digital Image Magnifying Based on Hidden Markov Model

Xingyi Zhong, Zhaojun Li and Xinyu Chen

Abstract In this paper we propose a new design technology for digital image magnifying based on Hidden Markov Model (HMM). First, we focus on the corresponding rules between state sequence and the matrix of observations, and establish the topological model for digital images in the basis of pseudo-two-dimensional structure. By introducing the estimation of related possibilities, we propose a parameter learning algorithm combined with data smoothing method, which is specially designed for classified digital image processing. To verify the algorithm's applicability and its actual result of image magnifying, we use Viterbi algorithm to Implement our newly proposed algorithm, and evaluate its performance in comparison with traditional interpolation methods. Compared to previous research, we expand the application category of HMM and provide new approaches for further research on image magnifying.

Keywords Image magnifying · Hidden Markov model · Pseudo-two-dimensional structure · Data smoothing processing · Viterbi algorithm

1 Introduction

After the proposition of image enlargement technology in 1984 by Tsai and Huang, image magnifying algorithms have been developed deeply, which can be divided into four categories. The first category is the interpolation algorithm,

X. Zhong (✉)
School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China
e-mail: 243559128@qq.com

Z. Li
School of Mathematics and Information Sciences, Guangzhou University, Guangzhou 510006, China

X. Chen
School of Civil Engineering, Guangzhou University, Guangzhou 510006, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_38

which requires the low resolution image align its position with the high resolution image, and then use the non-uniform interpolation technique to calculate the value of new pixel in high resolution images. Deconvolution calculation may be involved to improve the clarity of images. The second category is the algorithm based on frequency domain analysis. It estimates the frequency spectrum of high resolution image according to the relationship between spatial shift and phase shift, and then reconstruct the high resolution image with Frontier transform. The third type is reconstruction algorithm. The reconstruction algorithm constructs linear equations based on the influence of different resolution image pixel gray value, and obtains high resolution image by solving the linear equations. The fourth category is algorithm based on machine learning, which has obvious difference with the above three algorithms for image signal processing. This kind of algorithm pay more attention to the understanding of images' pixels content and data structure, therefore may usually obtained better performance than traditional method.

The Hidden Markov Model (HMM) has been successfully applied to acoustic signal modeling [1]. Nowadays, HMM has expanded its application range to the areas of image processing, speech synthesis, seismic exploration and bio-signal processing [1]. HMM is a random process composed by internal finite state Markov chain and a series of random functions [2], and one of the two stochastic processes is unpredictable, which can only be predicted by analyzing the observation sequence of another random process [1]. The below three basic problems need to be solved before applying HMM to image magnifying process [3]:

(1) Recognition
For a given observation sequence $O = (O_1, O_2, \ldots, O_T)$ representing the pixel value of images and a set of HMM parameters $\lambda = \{\pi, A, B\}$, The probability of model λ generating O can be denoted as $P(O|\lambda)$, which can be calculated by using Forward Backward Procedure [3] with a time Complexity of N^2T.
(2) The selection of best state chain
The Viterbi algorithm is usually used to determine the best state chain $Q = q_1 q_2 \ldots q_T$.
(3) The optimization of model parameters
Iterative method or gradient descend is usually used when adjusting the model parameters $\lambda = \{\pi, A, B\}$, aiming to get maximum value of $P(O|\lambda)$.

In the following sections, we establish a Hidden Markov model-based image magnifying algorithms, and analyze the application of HMM in image processing.

2 HMM-Based Image Magnifying Algorithm

2.1 Topological Structure of HMM

Define the main parameters of HMM involved in the magnifying algorithm of images as below:

(1) The state set of original image: $S = \{S_1, S_2, \ldots, S_N\}$, N denotes the number of state.
(2) The set of observing symbols: $V = \{v_1, v_2, \ldots, v_M\}$, M denotes the number of observing symbols.
(3) The model state q_t at time: $q_t \in S, 1 \le t \le T$, T is the length of observation sequence, The experienced state sequence is $Q = (q_1, q_2, \ldots, q_t)$.
(4) The probability distribution of State transition $A = (a_{ij})_{N \times N}$, the elements of A can be defined as $a_{ij} = P(q_{t+1} = S_j | q_t = S_i), 1 \le i, j \le N$.
(5) The initial state probability of model $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_N)$, $\boldsymbol{\pi}_i = P(q_1 = S_i), 1 \le i \le N$.
(6) The Observation set of model output $O = \{O_1, O_2, \ldots, O_M\}$, and observation value at time t is $O_t, O_t \in \{O_1, O_2, \ldots, O_M\}$.
(7) Observation probability matrix $B = (b_{jk})_{N \times M}$, where $b_{jk} = P(O_t = v_k | q_t = S_i)$, $1 \le k \le M, 1 \le i \le N$.

After selecting the main parameters of HMM, the model can be represented by parameters set $\boldsymbol{\lambda} = \{\boldsymbol{\pi}, A, B\}$ containing N states $(S_1, S_2, \ldots, S_N)$. In this paper, we proposed pseudo-two-dimensional structure for digital image model. Its status sequence is a first-order Markov chain. The elements of images' state set S is the gray value $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots, \mathbf{255}\}$, its topological structure is shown in Fig. 1.

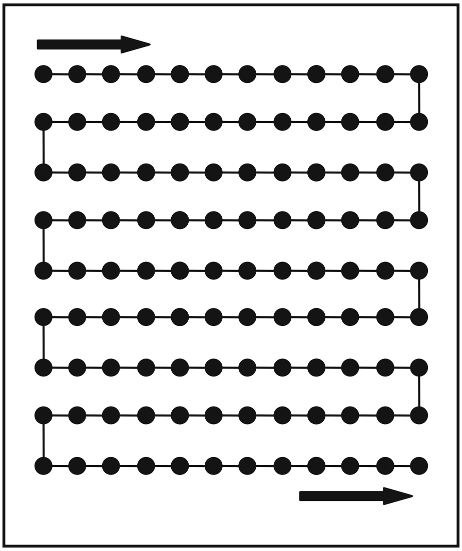


Fig. 1 Topological structure of digital image

According to Fig. 1, the gray value of each pixel is related with adjacent pixels, and a line of pixels are associated with the pixel line below, so that the link in digital image topological structure can be formed:

When $x \neq \mathbf{1}, M$ and y is even number, we have

$$P\left(q_{x,y}|q_{x-1,y}, q_{x,y-1}, q_{x-1,y-1}, \ldots, q_{r,s}, \ldots, q_{1,1}\right) = P\left(q_{x,y}|q_{x-1,y}\right) \tag{1}$$

When $x \neq \mathbf{1}, M$ and y is odd number, we have

$$P\left(q_{x,y}|q_{x-1,y}, q_{x,y-1}, q_{x-1,y-1}, \ldots, q_{r,s}, \ldots, q_{1,1}\right) = P\left(q_{x,y}|q_{x+1,y}\right) \tag{2}$$

When $x = \mathbf{1}, M$, we have

$$P\left(q_{x,y}|q_{x-1,y}, q_{x,y-1}, q_{x-1,y-1}, \ldots, q_{r,s}, \ldots, q_{1,1}\right) = P\left(q_{x,y}|q_{x,y-1}\right) \tag{3}$$

where $x = \mathbf{1}$ denotes the first row, $x = M$ represents the Mth row.

The transmission rule of digital image pixels from state sequence to observation sequence is shown in Fig. 2.

As shown in Fig. 2, The emission rule indicates that the grayscale value of every pixel (observation sequence) is relevant with corresponding pixel in state sequence.

$$P\left(O_{x,y}|q_{x,y}, \ldots, q_{1,1}, O_{x-1,y}, O_{x,y-1}, \ldots, O_{1,1}\right) = P\left(O_{x,y}|q_{x,y}\right) \tag{4}$$

This hypothesis we may greatly simplify the Hidden Markov Model for digital images.

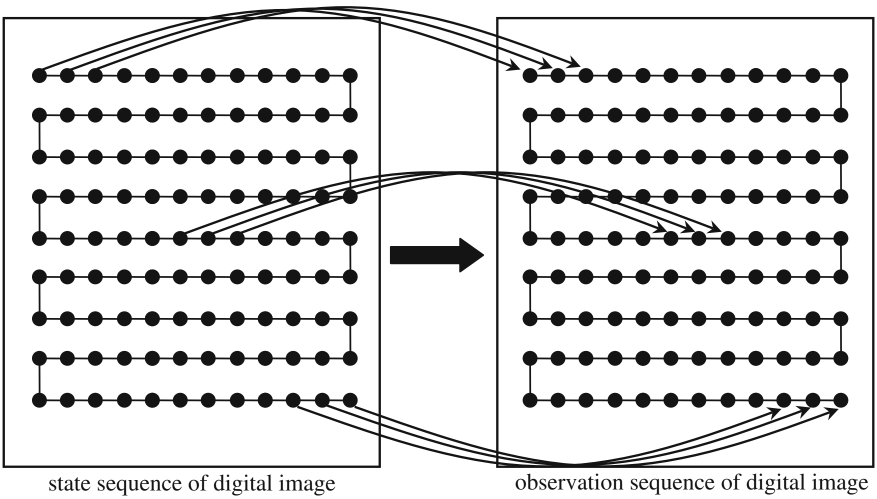


Fig. 2 Relation between state sequence and observation matrix

2.2 Parameter Learning Algorithm

The hidden Markov model used in speech recognition technology commonly use unsupervised learning algorithm, which may prone to get local minima. In this paper, we choose to use supervised learning algorithm with labeled data set, which can effectively avoid the problem of local minima. Suppose the given training data contains S observations sequences and corresponding state sequence $\{(O_1, I_1), (O_2, I_2), L, (O_s, I_s)\}$, According to Law of Large Numbers [4], we have $lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| < \varepsilon\right) = \mathbf{1}$, where S_n denotes the transmission frequency. When the training sample set is large enough, you can use frequency to estimate the probability.

(1) Parameter estimation

Suppose the sample position t is now at state i, and the frequency of position $t + \mathbf{1}$ transfer to state j is A_{ij}, then the estimation of state transition probability a_{ij} is

$$a_{ij} = \frac{A_{ij}}{\sum_{i=1}^{N} A_{ij}}, i = 1, 2, \ldots, N; j = 1, 2, \ldots, N \tag{5}$$

Suppose the sample state is j, and the frequency of observation k is B_{jk}, then the probability estimation $b_j(k)$ of observation $\mathbf{k}$ at state j is

$$b_j(k) = \frac{B_{jk}}{\sum_{j=1}^{M} B_{jk}}, j = 1, 2, \ldots, M; k = 1, 2, \ldots, M \tag{6}$$

Suppose the frequency of initial state of sample is q_i equals C_i, then the initial state probability estimation $\boldsymbol{\pi}_i$ is

$$\pi_i = \frac{C_i}{\sum_{i=0}^{255} C_i}, i = 0, 1, \ldots, 255 \tag{7}$$

(2) Data smoothing

In this paper, Laplace smoothing method [5] is adopted:

$$p_{Lap}(w_1, \ldots, w_n) = \frac{C(w_1, \ldots, w_n) + 1}{N + T} \tag{8}$$

where N represents the total number of training samples, T indicates the number of types of training samples.

(3) Classification for images

Since various types of unprocessed images have different features and different styles. There exist significant differences between various image types. As for portrait images, for example, the concentrated color blocks in the images are generally common and obvious, adjacent pixel values are almost the same, while the images of posters are usually created with messy color distribution and rich details. If we process the images with a single model without classification, the magnified effect will be poor. To improve the performance of the model training process and treatment effect on all kinds of images, we divide the images into several categories based on the content, such as scenery, cartoon, characters, posters and other categories. Then we train an HMM model for each category separately. After the training process of HMM model, appropriate type of model parameter for different types of images will be used, so that better effect can be reached [6].

2.3 Viterbi Algorithm

We choose Viterbi algorithm as the decoding approaches for HMM. The Viterbi algorithm is a dynamic deformation algorithm [3]. First define auxiliary variables $\boldsymbol{\psi}_t(i)$ and $\boldsymbol{\delta}_t(i)$. They satisfy the following relationship:

$$\delta_t(i) = \max_{q_{t-1}} P(q_t = S_i, O_1 O_2 \cdots O_t | \theta) \tag{9}$$

The main steps of Viterbi algorithm can be represented as follows:

(1) Initialization:

$$\delta_1(i) = \pi_i b_i(O_1), \psi_1(i) = 0, 1 \le i \le N \tag{10}$$

(2) Iteration:

$$\delta_t(j) = \max_{1 \le i \le N} \left[\delta_{t-1}(i) a_{ij}\right] b_j(O_t), 2 \le t \le T, 1 \le j \le N \tag{11}$$

$$\psi_t(j) = \operatorname{argmax}_{1 \le i \le N} \left[\delta_{t-1}(i) a_{ij}\right], 2 \le t \le T, 1 \le j \le N \tag{12}$$

(3) Finish:

$$q' = \max_{1 \le i \le N} [\delta_T(i)], q'_T = \operatorname{argmax}_{1 \le i \le N} [\delta_T(i)] \tag{13}$$

(4) Decoding (Solving the best state chain):

$$q'_t = \psi_{t+1}\left(q'_{t+1}\right), 1 \le t \le T - 1 \tag{14}$$

$P'(O/\lambda)$ and the best state chain $\left\{q'_1, q'_2, \ldots, q'_T\right\}$ can be solved by the Viterbi algorithm.

3 Simulation Test

We select a large number of digital images as the test samples to analyze the performance of the proposed algorithm. We write program with MATLAB according to the theory mentioned above, and design HMM algorithm to estimate model parameters. Parts of the result are shown in Fig. 3.

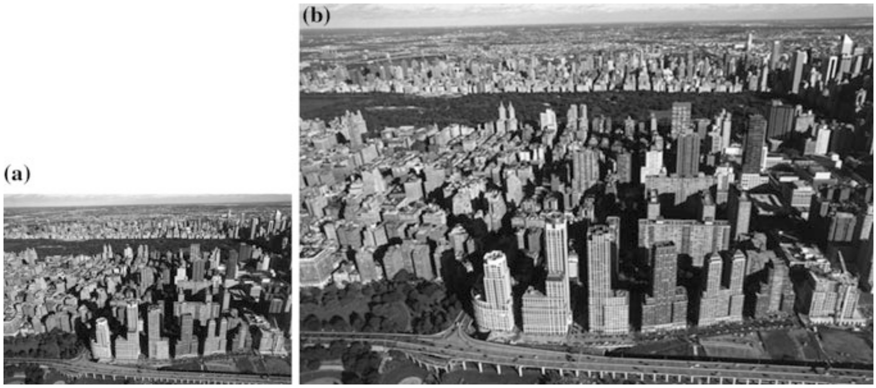


Fig. 3 Images magnifying result, **a** original image, **b** magnified image

By analyzing the processing result of the proposed algorithm, the magnified image has sharp edges and intact details, without any obvious fuzzy blocks or pseudo data points. Hence the images still maintain excellent visual effect after the magnifying process [7].

In order to evaluate the actual performance of the proposed algorithm objectively, we use the PSNR (peak signal to noise ratio) and NMSE (mean square error value after normalization) as the evaluation approaches for algorithm performance. The PSNR can be calculated as follow

$$PSNR = 10 \times lg\left(\frac{255^2}{MSE}\right) \tag{15}$$

$$MSE = \frac{\sum_{i=1}^{X}\sum_{j=1}^{Y}\left[P(i,j) - P'(i,j)\right]^2}{X \times Y} \tag{16}$$

$$NMSE = \frac{\sum_{i=1}^{X-1}\sum_{j=1}^{Y-1}[P(i,j) - P'(i,j)]^2}{\sum_{i=1}^{X-1}\sum_{j=1}^{Y-1}P(i,j)^2} \tag{17}$$

where **MSE** is Mean square error, **X**, **Y** is the horizontal and vertical length of the image, $\mathbf{P}(\mathbf{i},\mathbf{j})$, $\mathbf{P}'(\mathbf{i},\mathbf{j})$ is the grayscale value of the pixel in the original image and magnified image respectively.

To verify the effectiveness of the algorithm, we make a comparison experiment between the traditional bicubic interpolation algorithm, image amplification algorithms based on artificial neural network (ANN) and our newly proposed algorithm. Bicubic interpolation algorithm is a relatively simple image magnifying algorithm based on the traditional interpolation process, which is one of the most popular mainstream two-dimensional interpolation algorithms. Neural network and the new algorithm proposed in this paper are both learning algorithm based on the principle

Table 1 PSNR of three algorithms

	Image 1	Image 2	Image 3	Image 4
Proposed algorithm	45.136748	62.715503	50.598135	51.428452
Neutral network	45.024790	62.429747	50.490968	51.128536
Bicubic interpolation	45.136642	62.715433	50.598125	51.428446

Table 2 NMSE of three algorithms

Unit: $\mathbf{1 \times 10^3}$				
	Image 1	Image 2	Image 3	Image 4
Proposed algorithm	42.49134	6.02327	36.10733	22.99872
Neutral network	42.96973	6.19788	36.49637	23.69894
Bicubic interpolation	42.49179	6.02332	36.10737	22.99874

of image content understanding and summarizing [8]. Therefore they both have stronger capacity of details reconstruction and better treatment effect in comparison with other types of algorithm.

The original images in test sample set were selected from landscapes, figures, posters, building four categories, the performance test result are shown in Tables 1 and 2.

Apparently, the process effect of our newly proposed algorithm is obviously better than the neural network [9], which is an algorithm based on learning as the former one. When compared with traditional interpolation [10–12], a higher quality of reduction and smaller processing error can be seen in the proposed algorithm. The new algorithm's performance indicators are higher than bicubic interpolation algorithm in a variety of test environment. So the new algorithm based on hidden Markov model for image magnifying has satisfying performance and high practicality.

4 Conclusion

This article extends the application of Hidden Markov Models, establish the pseudo two dimensional topological structure and the state transition emission rules for digital images, and design a parameter learning algorithm and the Viterbi algorithm, which provides a new approach and technical support for image magnifying processing. By using MATLAB as a test environment, performance analysis has been done with actual digital images. The test data indicate that the image processing method based on hidden Markov model is more scientific, more reliable than other approaches. The new approaches can provide a new scientific method for image magnifying as an extensive application of Hidden Markov model.

References

1. Wang, Z.T., Cai, L.B.: Hidden Markov Models (HMM) and its application. J. Hunan Univ. Sci. Technol. **30**(4), 42–44 (2009)
2. Hu, Y.S., Zhu, A.E., Chen, G., Chen, Z.X.: A image classification algorithm based on two-dimensional hidden Markov model. J. Comput. Appl. **25**(4), 760–762 (2005)
3. Duan, H.M., Wang, J., Ma, L.H., Xu, R.: The application of hidden Markov model in speech recognition. J. Math. Technol. **18**(6), 16–20 (2002)
4. Mao, S.S., Cheng, Y.M., Pu, X.L.: Probability Theory and Mathematical Statistics, 2nd edn. Higher Education Press, Beijing (2011)
5. Laplace, P.S.: Philosophical Essay on Probabilities. Springer, New York (1995)
6. Zhu, M., Guo, C.S.: Hidden Markov model and its latest application and development. J. Comput. Syst. Appl. **19**(7), 255–259 (2010)
7. Liu, J.Y., Song, M.Q., Zhao, F., Yang, Z.Y.: Second order HMM based on context. J. Comput. Eng. **36**(10), 231–235 (2010)
8. Du, S.P.: The Baum-Welch algorithm of mixed order hidden Markov models. J. Yunnan Univ. (Nat. Sci.) **28**(2), 98–102 (2006)
9. Yang, D., Zhao, L.Z., Long, Z.: Examples of MATLAB Image Processing. Tsinghua University Press, Beijing (2013)
10. Gao, G.W.: Research on Some Problems of Denoising. Henan University, Zhengzhou (2008)
11. Shi, O.Y., Yang, J., Tian, X.: Identify CpG Island with MATLAB-based hidden Markov models. J. Comput. Appl. Softw. **25**(11), 214–216 (2008)
12. Zhou, P.: MATLAB Image Processing and Graphical User Interface Design. Tsinghua University Press, Beijing (2013)

An Evaluation Model in P2P File Sharing System

Pei-Hua Wang, Hai-Tao Lin and Xiao-Peng Yang

Abstract A peer-to-peer (P2P) file sharing system can be reduced into a system of addition-min fuzzy relation inequalities. For evaluating the satisfaction degree of the potential solutions in such system, we define a ranking function and establish the corresponding evaluation model. Numerical examples are provided to illustrate the feasibility and efficiency of the proposed evaluation model.

Keywords Evaluation model · P2P file sharing system · Fuzzy relation inequality · Satisfaction degree

1 Introduction

Fuzzy relation equation was introduced by Sanchez [1] for the first time. Since then theoretical method and practical application of fuzzy relation equation or inequality were investigated, with various compositions, e.g. *max-min*, *max-product* and *addition-min*. Associated with the research of fuzzy relation equation and inequality, the corresponding fuzzy relation mathematical programming problems were also studied. Scholars were interested in both linear [2–6] and nonlinear [7–10] fuzzy relation optimization problems.

P.-H. Wang
Guangzhou Vocational College of Science and Technology,
Guangzhou 510550, China

H.-T. Lin · X.-P. Yang (✉)
School of Mathematics and Statistics, Hanshan Normal University,
Chaozhou 521041, China
e-mail: happyyangxp@163.com

X.-P. Yang
School of Mathematics and Information Science, Guangzhou University,
Guangzhou 510006, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_39

Recently, J.-X. Li et al. pointed out that a P2P file sharing system could be reduced into a system of addition-min fuzzy relation inequalities [11]. The authors introduced some basic definitions on the addition-min fuzzy relation inequalities and proposed a novel method to find one of the minimal solutions. In order to avoid network congestion in the P2P file sharing system, S.-J. Yang [12] established a linear programming subject to addition-min fuzzy relation inequalities. Solution method based on pseudo-minimal index was employed to deal with the proposed problem. To improve the optimal management objective adopted in [12], X.-P. Yang proposed the max-min programming problem with addition-min fuzzy relation inequalities and developed a sub-problem method for obtaining the optimal solution.

In the paper, we aim to establish an evaluation model for evaluating the satisfaction degree of a potential solution in the P2P file sharing system. The remaining content is organized as follows. In Sect. 2 we present some necessary concepts and results about the P2P file sharing system. Evaluation model and method based on addition-min fuzzy relation inequalities are provided in Sect. 3. Section 4 gives some illustrative examples while Sect. 5 is simple conclusion.

2 Preliminaries

In this section we recall some relevant concepts and existing results about the P2P file sharing system. The readers can refer to [11–14] for more detail.

It has been expressed in [11] than the P2P file sharing system can be reduced into a system of addition-min fuzzy relation inequalities. There are n terminals (or users) in the P2P file sharing system, denoted by $A_1, A_2, \dots, A_n$. All the terminals in such system is connected. Each terminal is able to connect with and shares its local file to any other terminal. The j the terminal, i.e. A_j, sends file data to other terminals with quality level x_k, $j = 1, 2, \dots, n$. Suppose the quality requirement of download traffic of the ith terminal, i.e. A_i, is at least b_i, and the bandwidth between A_i and A_j is a_{ij}, $j = 1, 2, \dots, n$, $i = 1, 2, \dots, m$, $m \le n$. Then the P2P file sharing system can be described as

$$\begin{cases} a_{11} \wedge x_1 + a_{12} \wedge x_2 + \cdots + a_{1n} \wedge x_n \ge b_1, \\ a_{21} \wedge x_1 + a_{22} \wedge x_2 + \cdots + a_{2n} \wedge x_n \ge b_2, \\ \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\ a_{m1} \wedge x_1 + a_{m2} \wedge x_2 + \cdots + a_{mn} \wedge x_n \ge b_m. \end{cases} \tag{1}$$

After normalization, it is reasonable to assume that $a_{ij} \in [0, 1]$, $x_j \in [0, 1]$, $b_i \in (0, 1]$.

We call $X = [0, 1]^n$ the potential solution of the P2P file sharing system (or system (1)).

Let $I = \{1, 2, \dots, m\}$ and $J = \{1, 2, \dots, n\}$ be two index sets, then system (1) can be written as

$$\sum_{j \in J} a_{ij} \wedge x_j \ge b_i, \quad \forall i \in I, \tag{2}$$

or

$$A \odot x^T \geq b^T, \tag{3}$$

where $A = (a_{ij})_{m\times n}$, $x = (x_1, x_2, \ldots, x_n)$, $b = (b_1, b_2, \ldots, b_m)$, and

$$\begin{aligned}&(a_{i1}, a_{i2}, \ldots, a_{in}) \odot (x_1, x_2, \ldots, x_n)^T \\ &= a_{i1} \wedge x_1 + a_{i2} \wedge x_2 + \cdots + a_{in} \wedge x_{1n}.\end{aligned}$$

Definition 1 (*See* [11]) Let $x^1 = (x_1^1, x_2^1, \ldots, x_n^1)$, $x^2 = (x_1^2, x_2^2, \ldots, x_n^2) \in X$, we define:
(i) $x^1 \leq x^2$ if $x_j^1 \leq x_j^2$, $\forall j \in J$;
(ii) $x^1 < x^2$ if $x^1 \leq x^2$ and there are some $j \in J$ such that $x_j^1 < x_j^2$.

In what follows we shall denote the dual of order relation '$<$' and '$\leq$' by the symbol '$>$' and '$\geq$', respectively. Obviously, the operator '$\leq$' forms a partial order relation on X and $(X, \leq)$ becomes a lattice.

We denote the solution set of system (1) by $X(A, b) = \{x \in X | A \odot x^T \geq b^T\}$.

Definition 2 (*See* [14]) A solution $\hat{x} \in X(A, b)$ is said to be the maximum (or greatest) solution of system (1) if and only if $x \leq \hat{x}$ for all $x \in X(A, b)$. A solution $\check{x} \in X(A, b)$ is said to be a lower (or minimal) solution of system (1) if and only if $x \leq \check{x}$ implies $x = \check{x}$ for any $x \in X(A, b)$. A solution $\dot{\check{x}} \in X(A, b)$ is said to be the minimum solution of system (1) if and only if $x \geq \dot{\check{x}}$ for all $x \in X(A, b)$.

Definition 3 (*See* [14]) System (1) is said to be consistent if $X(A, b) \neq \emptyset$. Otherwise, it is said to be inconsistent.

Obviously, when system (1) is consistent, $\hat{x} = (1, 1, \ldots, 1)$ is the maximum solution. As shown in [12], if $X(A, b) \neq \emptyset$, then $X(A, b) = \bigcup_{\check{x} \in \check{X}(A,b)} \{x | \check{x} \leq x \leq \hat{x}\}$, where $\check{X}(A, b)$ is the set of all lower solutions of system (1). Now we introduce some properties and existing results on system (1).

Theorem 1 (See [12, 14]) *For system (1), we have:*

- *(i) Equation (1) is consistent if and only if* $\sum_{j\in J} a_{ij} \geq b_i$ *for arbitrary* $i \in I$*;*
- *(ii) Let* $x^* \in X(A, b)$, $x \in X$. $x^* \leq x$ *implies* $x \in X(A, b)$*;*
- *(iii) Let* $x', x \in X$ *and* $x \leq x'$. $x' \notin X(A, b)$ *implies* $x \notin X(A, b)$*;*
- *(iv) Let* $x \in X(A, b)$*, if* $\sum_{j\in J} a_{ij} = b_i$ *for some* $i \in I$*, then* $(a_{i1}, a_{i2}, \ldots, a_{in}) \leq x$.

Theorem 2 (See [12, 13]) *Let* $x \in X(A, b)$ *be a solution of system (1), then we have:*

- *(i)* $x > 0$*;*
- *(ii) For arbitrary* $i \in I$, $j \in J$,

$$x_j \geq b_i - \sum_{k\in J-\{j\}} a_{ik} \wedge x_k \geq b_i - \sum_{k\in J-\{j\}} a_{ik};$$

(iii) For arbitrary $i \in I$, $j \in J$,

$$a_{ij} \geq b_i - \sum_{k \in J - \{j\}} a_{ik} \wedge x_k \geq b_i - \sum_{k \in J - \{j\}} a_{ik}.$$

Let $\dot{\check{x}} = (\dot{\check{x}}_1, \dot{\check{x}}_2, \ldots, \dot{\check{x}}_n)$, where

$$\dot{\check{x}}_j = \max_{i \in I} \{0, b_i - \sum_{k \in J - \{j\}} a_{ik}\}, \tag{4}$$

$j = 1, 2, \ldots, n$. Then the uniqueness of the lower solution of system (1) can be checked by the following theorem.

Theorem 3 (See [12]) *System (1) has the unique lower solution if and only if $\dot{\check{x}}$ is a solution of (1), i.e. $\dot{\check{x}} \in X(A, b)$. In particular, when (1) has the unique lower solution, $\dot{\check{x}}$ is the unique lower solution of system (1).*

Corollary 1 (See [14]) *$\dot{\check{x}}$ is the minimum solution if and only if $\dot{\check{x}}$ is a solution of system (1).*

From Theorem 3 and Corollary 1, we know that $\dot{\check{x}}$ is the potential minimum solution of system (1). The minimum solution doesn't always exist, but if it does so, it must be $\dot{\check{x}}$.

3 Evaluation Model Based on Addition-Min Fuzzy Relation Inequalities

For any $x \in X$, we can define its feasible index set in system (1) as follows:

$$I_x = \{i \in I | a_{i1} \wedge x_1 + a_{i2} \wedge x_2 + \cdots + a_{in} \wedge x_n \geq b_i\}. \tag{5}$$

In the P2P file sharing system, $i \in I_x$ indicates the download requirement of the ith terminal is satisfied. Based on the consideration of the requirements of terminals, for arbitrary potential solution $x, y \in X$, x is considered to be better than y if x satisfies more terminals with regard to the download requirement.

Definition 4 Let $x \in X$ be an arbitrary potential solution of the P2P file sharing system. Then

(i) x is said to be a completely satisfactory solution of the P2P file sharing system, if x satisfies all of the constraint inequalities in system (1), i.e. $I_x = I$;
(ii) x is said to be a dissatisfactory solution of the P2P file sharing system, if x satisfies non of the constraint inequalities in system (1), i.e. $I_x = \emptyset$;

(iii) x is said to be a partly satisfactory solution of the P2P file sharing system, if x is neither a completely satisfactory solution, nor a dissatisfactory solution, i.e. x satisfies some, but not all, of the constraint inequalities in system (1), i.e. $I_x \neq \emptyset$ and $I_x \neq I$.

Definition 5 Let $x, y \in X$ be two potential solutions of the P2P file sharing system. x is said to be superior to (more satisfactory than) y, if $|I_x| > |I_y|$. x is said to be equivalent to y, if $|I_x| = |I_y|$.

Theorem 4 *Let $x \in X$. Then x is an completely satisfactory solution of the P2P file sharing system if and only if $x \in X(A, b)$*

Proof Since $x \in X$, we have

$$\begin{aligned} & x \text{ is an completely satisfactory solution of the P2P file sharing system} \\ & \Leftrightarrow I_x = I \\ & \Leftrightarrow \text{For any} i \in I, a_{i1} \wedge x_1 + a_{i2} \wedge x_2 + \cdots + a_{in} \wedge x_n \geq b_i \\ & \Leftrightarrow x \in X(A, b). \end{aligned} \tag{6}$$

The proof is complete. □

Corollary 2 *Let $x \in X$ and system (1) be consistent with minimum solution $\dot{\check{x}}$, i.e. $A \odot \dot{\check{x}}^T \geq b^T$. Then x is an completely satisfactory solution of the P2P file sharing system if and only if $x \geq \dot{\check{x}}$.*

Proof Notice that $\dot{\check{x}}$ is the minimum solution of system (1), while $\hat{x} = (1, 1, \ldots, 1)$ is the maximum solution. On the other hand, $x \in X = [0, 1]^n$. Hence $x \in X(A, b)$ if and only if $x \geq \dot{\check{x}}$. The rest of the proof is due to Theorem 4. □

Definition 6 Let p, q be two real number. Define the signal function as follows:

$$Sign(p, q) = \begin{cases} 1, & \text{if } p \geq q; \\ 0, & \text{if } p < q. \end{cases} \tag{7}$$

Let $x \in X$. $A = (a_{ij})$, $A_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, n$. We define the following ranking function for evaluating the satisfaction degree of a potential solution of the P2P file sharing system.

$$f(x) = \sum_{i=1}^{m} \frac{Sign(A_i \odot x^T, b_i)}{m} \tag{8}$$

Theorem 5 *For any any potential solution $x \in X$ in the P2P file system, it holds that $0 \leq f(x) \leq 1$. Furthermore,*

(i) x is a completely satisfactory solution if and only if $f(x) = 1$*;*
(ii) x is a partly satisfactory solution if and only if $0 < f(x) < 1$*;*
(iii) x is a dissatisfactory solution if and only if $f(x) = 0$.

Proof According to (7), for arbitrary $i \in I$, $Sign(A_i \odot x^T, b_i) \in \{0, 1\}$. This indicates

$$0 \leq \sum_{i=1}^{m} Sign(A_i \odot x^T, b_i) \leq m.$$

Notice that

$$f(x) = \sum_{i=1}^{m} \frac{Sign(A_i \odot x^T, b_i)}{m} = \frac{\sum_{i=1}^{m} Sign(A_i \odot x^T, b_i)}{m}.$$

Hence

$$0 \leq f(x) \leq 1.$$

(i) Considering $Sign(A_i \odot x^T, b_i) \in \{0, 1\}$, $\forall i \in I$, it is easy to check that

$$\begin{aligned} f(x) = 1 &\Leftrightarrow \frac{\sum_{i=1}^{m} Sign(A_i \odot x^T, b_i)}{m} = 1 \\ &\Leftrightarrow \sum_{i=1}^{m} Sign(A_i \odot x^T, b_i) = m \\ &\Leftrightarrow Sign(A_i \odot x^T, b_i) = 1,\ \forall i \in I \\ &\Leftrightarrow A_i \odot x^T \geq b_i,\ \forall i \in I \\ &\Leftrightarrow A \odot x^T \geq b^T \\ &\Leftrightarrow x \in X(A, b), \text{i.e. } x \text{ is a completely satisfactory solution.} \end{aligned} \tag{9}$$

(ii) Similar to (i), we get

$$\begin{aligned} 0 < f(x) < 1 &\Leftrightarrow 0 < \frac{\sum_{i=1}^{m} Sign(A_i \odot x^T, b_i)}{m} < 1 \\ &\Leftrightarrow 0 < \sum_{i=1}^{m} Sign(A_i \odot x^T, b_i) < m \\ &\Leftrightarrow Sign(A_{i'} \odot x^T, b_{i'}) = 1 \text{ for some } i' \in I, \\ &\quad\ \text{and } Sign(A_{i''} \odot x^T, b_{i''}) = 0 \text{ for some } i'' \in I \\ &\Leftrightarrow A_i \odot x^T \geq b_i \text{ for some } i' \in I, \end{aligned} \tag{10}$$

$$\text{and } A_i \odot x^T < b_i \text{ for some } i'' \in I$$
$$\Leftrightarrow I_x \neq \emptyset \text{ and } I_x \neq I$$
$$\Leftrightarrow x \text{ is a partly satisfactory solution.}$$

(iii) Similarly,

$$\begin{aligned} f(x) = 0 &\Leftrightarrow \frac{\sum_{i=1}^{m} Sign(A_i \odot x^T, b_i)}{m} = 0 \\ &\Leftrightarrow Sign(A_i \odot x^T, b_i) = 0, \ \forall i \in I \\ &\Leftrightarrow A_i \odot x^T < b_i, \ \forall i \in I \\ &\Leftrightarrow I_x = \emptyset \\ &\Leftrightarrow x \text{ is a dissatisfactory solution.} \end{aligned} \tag{11}$$

□

Theorem 6 *For any $x, y \in X$, x is superior to y if and only if and only if $f(x) > f(y)$, while x is equivalent to y if and only if and only if $f(x) = f(y)$. In particular, x is superior or equivalent to y if and only if and only if $f(x) \geq f(y)$.*

Proof According to (5),

$$i \in I_x \Leftrightarrow A_i \odot x^T \geq b_i \Leftrightarrow Sign(A_i \odot x^T, b_i) = 1,$$

and conversely,

$$i \notin I_x \Leftrightarrow A_i \odot x^T < b_i \Leftrightarrow Sign(A_i \odot x^T, b_i) = 0.$$

Furthermore we have

$$\begin{aligned} |I_x| &= |I_x| + 0 \\ &= \sum_{i \in I_x} 1 + \sum_{i \notin I_x} 0 \\ &= \sum_{i \in I_x} Sign(A_i \odot x^T, b_i) + \sum_{i \notin I_x} Sign(A_i \odot x^T, b_i) \\ &= \sum_{i \in I} Sign(A_i \odot x^T, b_i) \\ &= m \cdot f(x). \end{aligned} \tag{12}$$

In the same way we can verify that $|I_y| = m \cdot f(y)$. Hence it is obvious that $|I_x| > |I_y|$ (or $|I_x| = |I_y|$, $|I_x| \geq |I_y|$) if and only if $f(x) > f(y)$ (or $f(x) = f(y)$, $f(x) \geq f(y)$), and the rest of the proof can be easily obtained following Definition 5. □

Corollary 3 *Suppose x, y, z are arbitrary completely satisfactory solution, partly satisfactory solution and dissatisfactory solution, respectively. Then it is clear that x is superior to y and y is superior to z.*

4 Numerical Examples

Example 1 A six-users BitTorrent-like Peer-to-Peer file sharing system is reduced into the following addition-min fuzzy relation inequalities:

$$A \odot x^T \geq b^T, \tag{13}$$

where

$$A = (a_{ij}) = \begin{bmatrix} 0 & 0.6 & 0.8 & 0.5 & 0.6 & 0.9 \\ 0.5 & 0 & 0.7 & 0.9 & 0.8 & 0.5 \\ 0.8 & 0.7 & 0 & 0.4 & 0.7 & 0.8 \\ 0.7 & 0.5 & 0.6 & 0 & 0.8 & 0.6 \\ 0.8 & 0.6 & 0.9 & 0.7 & 0 & 0.7 \\ 0.6 & 0.7 & 0.7 & 0.9 & 0.5 & 0 \end{bmatrix},$$

$$b = (b_1, b_2, \dots, b_6) = (2.8, 3.0, 2.9, 2.5, 3.2, 3.0),$$

$$x = (x_1, x_2, \dots, x_6) \in [0, 1]^6,$$

and $\odot$ is the addition-min composition. Here, a_{ij} represents the bandwidth between ith user and jth user, x_j is the quality level on which the file data are sent from jth user, b_i is the quality requirement of download traffic of ith user. Take three potential solution of system (13) as follows: $x^1 = (0.3, 0.3, 0.4, 0.5, 0.4, 0.3)$, $x^2 = (0.8, 0.6, 0.7, 0.7, 0.9, 0.7)$ and $x^3 = (0.6, 0.6, 0.6, 0.6, 0.6, 0.6)$. We aim to determine whether $x^j, j = 1, 2, 3$, is a completely satisfactory solution, a partly satisfactory solution, or a dissatisfactory solution of the corresponding P2P file sharing system.

Solution
Since there are 6 constraint inequalities in system (13), let the index set $I = \{1, 2, \dots, 6\}$. Denote $A_i = (a_{i1}, a_{i2}, \dots, a_{i6})$, $i = 1, 2, \dots, 6$. For making the decision to evaluation of x^k, we compute I_{x^k} by (5) in the following, $k = 1, 2, 3$.

For $x^1 = (0.3, 0.3, 0.4, 0.5, 0.4, 0.3)$,

$$A_1 \odot {x^1}^T = 0 \wedge 0.3 + 0.6 \wedge 0.3 + 0.8 \wedge 0.4 + 0.5 \wedge 0.5 + 0.6 \wedge 0.4 + 0.9 \wedge 0.3 = 1.9 < 2.8 = b_1. \tag{14}$$

Similarly we get

$$A_2 \odot {x^1}^T = 0.5 \wedge 0.3 + 0 \wedge 0.3 + 0.7 \wedge 0.4 + 0.9 \wedge 0.5 + 0.8 \wedge 0.4 + 0.5 \wedge 0.3$$
$$= 1.9 < 3.0 = b_2,$$
$$A_3 \odot {x^1}^T = 0.8 \wedge 0.3 + 0.7 \wedge 0.3 + 0 \wedge 0.4 + 0.4 \wedge 0.5 + 0.7 \wedge 0.4 + 0.8 \wedge 0.3$$

$$
\begin{aligned}
&= 1.7 < 2.9 = b_3,\\
A_4 \odot x^{1^T} &= 0.7 \wedge 0.3 + 0.5 \wedge 0.3 + 0.6 \wedge 0.4 + 0 \wedge 0.5 + 0.8 \wedge 0.4 + 0.6 \wedge 0.3\\
&= 1.7 < 2.5 = b_4,\\
A_5 \odot x^{1^T} &= 0.8 \wedge 0.3 + 0.6 \wedge 0.3 + 0.9 \wedge 0.4 + 0.7 \wedge 0.5 + 0 \wedge 0.4 + 0.7 \wedge 0.3\\
&= 1.8 < 3.2 = b_5,\\
A_6 \odot x^{1^T} &= 0.6 \wedge 0.3 + 0.7 \wedge 0.3 + 0.7 \wedge 0.4 + 0.9 \wedge 0.5 + 0.5 \wedge 0.4 + 0 \wedge 0.3\\
&= 1.9 < 3.0 = b_6,
\end{aligned}
\tag{15}
$$

Hence $I_{x^1} = \emptyset$ and x^1 is a dissatisfactory solution.

In the same say, for $x^2 = (0.8, 0.6, 0.7, 0.7, 0.9, 0.7)$, we can compute

$$
\begin{aligned}
A_1 \odot x^{2^T} &= 3.1 \geq 2.8 = b_1,\\
A_2 \odot x^{2^T} &= 3.2 \geq 3.0 = b_2,\\
A_3 \odot x^{2^T} &= 3.2 \geq 2.9 = b_3,\\
A_4 \odot x^{2^T} &= 3.2 \geq 2.5 = b_4,\\
A_5 \odot x^{2^T} &= 3.5 \geq 3.2 = b_5,\\
A_6 \odot x^{2^T} &= 3.1 \geq 3.0 = b_6,
\end{aligned}
\tag{16}
$$

Hence $I_{x^2} = I$ and x^2 is a completely satisfactory solution.

For $x^3 = (0.8, 0.6, 0.7, 0.7, 0.9, 0.7)$, we compute

$$
\begin{aligned}
A_1 \odot x^{3^T} &= 2.9 \geq 2.8 = b_1,\\
A_2 \odot x^{3^T} &= 2.8 < 3.0 = b_2,\\
A_3 \odot x^{3^T} &= 2.8 < 2.9 = b_3,\\
A_4 \odot x^{3^T} &= 2.9 \geq 2.5 = b_4,\\
A_5 \odot x^{3^T} &= 3.0 < 3.2 = b_5,\\
A_6 \odot x^{3^T} &= 2.9 < 3.0 = b_6,
\end{aligned}
\tag{17}
$$

It is clear that $I_{x^3} = \{1, 4\}$. Hence $I_{x^3} \neq \emptyset$ and $I_{x^3} \neq I$, x^3 is a partly satisfactory solution.

Example 2 A BitTorrent-like Peer-to-Peer file sharing system is reduced into the following system of addition-min fuzzy relation inequalities:

$$\begin{pmatrix} 0.8 & 0.3 & 0.5 & 0.2 & 0.1 & 0.3 \\ 0.5 & 0.2 & 0.3 & 0.4 & 0.7 & 0.2 \\ 0.3 & 0.4 & 0.2 & 0.7 & 0.3 & 0.4 \\ 0.3 & 0.9 & 0.8 & 0.5 & 0.8 & 0.1 \\ 0.9 & 0.5 & 0.1 & 0.7 & 0.8 & 0.2 \end{pmatrix} \odot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \geq \begin{pmatrix} 1.5 \\ 2.0 \\ 1.8 \\ 2.9 \\ 2.6 \end{pmatrix}. \tag{18}$$

Compare the satisfaction degree of the potential solutions below: $x^1 = (0.8, 0.8, 0.6, 0.5, 0.4, 0.2)$, $x^2 = (0.6, 0.3, 0.1, 0.2, 0.4, 0.3)$, $x^3 = (0.7, 0.9, 0.6, 0.5, 0.6, 0)$, $x^4 = (0.7, 0.9, 0.6, 0.5, 0.6, 0.1)$.

Solution

Suppose the matrix form of system (18) is $A \odot x^T \geq b^T$ and the row vectors of A is $A_1, A_2, \ldots, A_5$.

For x^1, $A_1 \odot x^1 = 0.8 \wedge 0.8 + 0.3 \wedge 0.8 + 0.5 \wedge 0.6 + 0.2 \wedge 0.5 + 0.1 \wedge 0.4 + 0.3 \wedge 0.2 = 2.1$. Similarly, $A_2 \odot x^1 = 2.0$, $A_3 \odot x^1 = 1.9$, $A_4 \odot x^1 = 2.7$, $A_5 \odot x^1 = 2.5$. Thus

$$\begin{cases} Sign(A_1 \odot x^1, b_1) = Sign(2.1, 1.5) = 1, \\ Sign(A_2 \odot x^1, b_2) = Sign(2.0, 2.0) = 1, \\ Sign(A_3 \odot x^1, b_3) = Sign(1.9, 1.8) = 1, \\ Sign(A_4 \odot x^1, b_4) = Sign(2.7, 2.9) = 0, \\ Sign(A_5 \odot x^1, b_5) = Sign(2.5, 2.6) = 0. \end{cases} \tag{19}$$

Furthermore we have $f(x^1) = \frac{\sum_{i=1}^{5} Sign(A_i \odot x^1, b_i)}{5} = 0.6$. In the same way we can compute $f(x^2) = 0.2$ and $f(x^3) = 0.6$, $f(x^4) = 0.8$. It is obvious that

$$0 < f(x^2) \leq f(x^1) = f(x^3) \leq f(x^4) < 1. \tag{20}$$

According to Theorems 5 and 6, all these four potential solutions are partly satisfactory solution. In addition, it follows from (20) that x^4 is superior to x^1, x^2 and x^3, both x^1 and x^3 are superior to x^2, x^1 is equivalent to x^3.

5 Conclusion

Addition-min fuzzy relation inequality was applied to characterize the P2P file sharing system [11]. For evaluating the satisfaction degree of potential solution of the P2P file sharing system, some relevant properties of addition-min fuzzy relation inequality are studied and the ranking function is defined. Based on the ranking function, we establish an evaluation model to the P2P file sharing system with two illustrative examples.

Acknowledgments Thanks to the support by the Innovation and Building Strong School Project of Colleges of Guangdong Province (2015KQNCX094)

References

1. Sanchez, E.: Resolution of composite fuzzy relation equations. Inf. Control **30**, 38–48 (1976)
2. Fang, S.-C., Li, G.: Solving fuzzy relation equations with a linear objective function. Fuzzy Sets Syst. **103**, 107–113 (1999)
3. Loetamonphong, J., Fang, S.-C.: Optimization of fuzzy relation equations with max-product composition. Fuzzy Sets Syst. **118**, 509–517 (2001)
4. Wu, Y.-K., Guu, S.-M.: Minimizing a linear function under a fuzzy max-min relational equation constraint. Fuzzy Sets Syst. **150**, 147–162 (2005)
5. Qu, X.-B., Wang, X.-P.: Minimization of linear objective functions under the constraints expressed by a system of fuzzy relation equations. Inf. Sci. **178**, 3482–3490 (2008)
6. Chang, C.-W., Shieh, B.-S.: Linear optimization problem constrained by fuzzy max-min relation equations. Inf. Sci. **234**, 71–79 (2013)
7. Lu, J., Fang, S.-C.: Solving nonlinear optimization problems with fuzzy relation equations constraints. Fuzzy Sets Syst. **119**, 1–20 (2001)
8. Molai, A.A.: A new algorithm for resolution of the quadratic programming problem with fuzzy relation inequality constraints. Comput. Ind. Eng. **72**, 306–314 (2014)
9. Shivanian, E., Khorram, E.: Monomial geometric programming with fuzzy relation inequality constraints with max-product composition. Comput. Ind. Eng. **56**, 1386–1392 (2009)
10. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.: Single-variable term semi-latticized fuzzy relation geometric programming with max-product operator. Inf. Sci. **325**, 271–287 (2015)
11. Li, J.-X., Yang, S.-J.: Fuzzy relation equalities about the data transmission mechanism in bittorrent-like peer-to-peer file sharing systems. In: Proceedings of the 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2012, pp. 452–456
12. Yang, S.-J.: An algorithm for minimizing a linear objective function subject to the fuzzy relation inequalities with addition-min composition. Fuzzy Sets Syst. **255**, 41–51 (2014)
13. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.:Multi-level linear programming subject to addition-min fuzzy relation inequalities with application in peer-to-peer file sharing system. J. Intell. Fuzzy Syst. **28** 2679–2689 (2015)
14. Yang, X.-P., Zhou, X.-G., Cao, B.-Y.: Min-max programming problem subject to addition-min fuzzy relation inequalities. IEEE Trans. Fuzzy Syst. **24**(1), 111–119 (2016)

Part IV
Factor Space and Factorial Neural Networks

Factor Space and Normal Context on Concept Description

Kai-qi Zou and Li-na Shi

Abstract Normal concept analysis is an important mathematical branch on knowledge representation, but there are some defect points blocking the development of the theory. The authors of the paper suggests that the involution principle is put on basic concepts only and we does not need to emphasize the bi-directional involution. The single-directional involution is enough.

Keywords Factor space · Context concept analysis · Concept lattice · Basic concept semi-lattice · Involution

1 Introduction

Normal concept analysis is an important mathematical branch on knowledge representation initialed on 1982 by Wille [1].

Definition 1 [2] A formal context is a triple $K = (G, M, I)$, where G is a group of objects, M a set of attributes and $I \subseteq G \times M$ a relation, $(u, a) \in I$ means that the object u has the attribute a.

For any $A \subseteq G$, denote

$$f(A) = a \in M \mid \forall u \in A; (u, a) \in I,$$

which is the set of common attributes of objects in A; For any $B \subseteq G$, denote

$$g(B) = u \in G \mid \forall a \in B; (u, a) \in I,$$

which is the set of objects having attributes in B.

K. Zou (✉) · L. Shi
Dalian University, Dalian 116622, China
e-mail: zoukq@vip.sina.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_40

Definition 2 [2] Given a formal context $K = (G, M, I)$, $\alpha = (A, B)$ is called a *concept* with respect to K if

$$f(A) = B \text{ and } g(B) = A. \quad (1)$$

where A and B are called the *extension* and *intension* of α respectively. The condition (1) can be called the *involution*.

The most important contribution of sets theory is that it opens the door of mathematics towards to knowledge representation: Any extension of a concept is a set. But, is any set an extension of a concept? While answers No! A set may not be an extension of a concept except the condition of involution is satisfied. This is the most important contribution of Wille in his formal concept analysis.

Given a formal context, there are a lot of concepts forming a lattice, called a concept lattice by Wille. The main topics in book [2] are focused on how to generate several kinds of concept lattices.

Apart from the successful works of formal concept analysis, we have to ask some questions on Definition 2 in Sect. 2: Is a concept really satisfy the involution condition? Is there exist a concept lattice according to (AND, OR) intension operation pair? The answers given in the Sect. 2 will be: Only basic concepts satisfy involution condition. Basic concepts with respect to a formal context K form a semi-lattice with respect to the operation AND. The authors of the paper does not belittle Wille's creation, but attributes a little suggestion to his theory. Comparing formal concept analysis with factor space on the point, we have a brief commentary left in Sect. 3. Conclusion is left in Sect. 4.

2 Vulnerabilities in Formal Concept Analysis

Concept is the unite of thinking, and can be, in common sense, described from two sides, intension and extension, simultaneously: Intension indicates attribute and extension identifies object. Given two concepts α_i with intension a_i and extension $E_i (i = 1, 2)$, we call α_1 is a sub-concept of α_2 denoted as $\alpha_1 \leq \alpha_2$ if a_1 implies a_2 (or $E_1 \subseteq E_2$ equivalently). For a set of concepts C, $(C, \leq)$ is a partially ordered set. There are three logical operations can be defined on C: 'AND, OR and NON-'. They can be defined by $\wedge$, $\vee$ and $\sim$ on intensions or by $\cap$, $\cup$ and ^{c} on extensions respectively. For needs, we put a temporary definition as follows:

Definition 3 Given a partially ordered concept set $C = (C, \leq)$, $(C, \wedge)$ is called a *concept semi-lattice*. If C is closed for the operation $\wedge$, while $\wedge = \inf$, the greatest lower bound with respect to $\leq$. If $(C, \wedge)$ is a semi-lattice, then each concepts in C is called a *basic concept* with respect to $(C, \wedge)$.

Is a concept really satisfy the involution? Yes, if it is a basic concept with respect to a concept semi-lattice. Unfortunately, it is not true for general concepts. For example, the concept 'Enlisted soldier '=' Male enlisted soldier 'OR' Female enlisted soldier'. The two concepts in right side are basic concepts and receive involution both, but 'Enlisted soldier' is not a basic concept, it does not receive involution! General speaking, if two basic concepts α_1 and α_2, neither $\alpha_1 \leq \alpha_2$ nor $\alpha_2 \leq \alpha_1$ are given, then $\alpha = \alpha_1$ OR α_2 does not obey involution.

Therefore, it is better that we should have a pre-modifier 'Basic' added to concept in Definition 2. Since basic concepts are only closed under operation AND but not OR, all basic concepts do not form a basic concept lattice but a semi-lattice.

3 Concept Description by Factor Space Theory

Factor space was initiated by Wang [3], the same time as that of Wille. Factor space aims to provide a general framework of things and thinking process. It has published a lot of articles and books before, we only cite a few of Refs. [4–10], and Wang focuses on data science recently [11–13]. Factor space was defined as follows:

Definition 4 [4] A factor space defined on universe of discussion U is a family of sets $\psi = (\{X(f)\}_{(f \in F)}; U)$ satisfying

(1) $F = (F, \vee, \wedge, ^c, \mathbf{1}, \mathbf{0})$ is a complete Boolean algebra
(2) $X(0) = \Phi$;
(3) For any $T \subseteq F$, if $f \mid f \in T$ are *irreducible* (i.e., $s \neq t \Rightarrow s \wedge t = \mathbf{0}(s, t \in T)$), then

$$X(\{f \mid f \in T\}) = \prod_{f \in T} X(f)(\prod \text{stands for Cartesianproduct});$$

(4) $\forall f \in T$, there is a mapping with same symbol $f : f \to X(f)$.

$f \in F$ is called a factor on U. $X(f)$ is called the *state space* of factor f.

Factors $f_1, \ldots, f_n$ map each object to its state in $X(f_1) \times \cdots \times X(f_n)$. The Universe of discussion U is mapped into $G = \mathbf{a} = (a_1, \ldots, a_n) \mid \exists u \in U; a_1 = f_1(u), \ldots, a_n = f_n(u)$, which is called the *background relation*. Any $\mathbf{a} = (a_1, \ldots, a_n) \in G$ is called a (real) *quality-atom*, which is a sentence "$a_1 = f_1(u)$ *and* … *and* $a_n = f_n(u)$", where $a_j = f_j(u)$ is reading as "the state is a_j with respect to f_j". For any $\mathbf{a} = (a1, \ldots, a_n) \in G, \mathbf{a} = (a_1, \ldots, a_n)$ is called a *real quality-atom*, which is the unite of quality description. Denote $P(G) = B \mid B \in G$, which corresponds to C, the set of all concepts generated by those quality atoms. It is obvious that $(C, \vee, \wedge, \sim)$ is isomorphism to $(P(B), \cup, \cap, ^c)$ and called the *Boolean concept algebra on* ψ.

Consider the inversed mapping of F on G

$$F^{-1} : G \to P(U)$$
$$\mathbf{a} = (a_1, \ldots, a_n) \mid [\mathbf{a}] = \{u \in U \mid F(u) = \mathbf{a}\}.$$

Extend the domain of mapping F^{-1} from G to its power $P(G)$ again:

$$F^{-1} : P(G) \to P(U)$$
$$\vee a, a', \ldots, a^{\wedge} \mid [a], [a'], \ldots, [a^{\wedge}].$$

Proposition 1 *The extended mapping F^{-1} is isomorphism between ($P(G)$, ∨, ∧, ∼) and ($P(U)$, ∪, ∩, ^{c}).*

Definition 5 [14] For any $A \in G$, denote $E = F^{-1}(A)$, we call $\alpha = (E, A)$ a concept with intension A and extension E. The set of all concepts $\mathfrak{J} = (\mathfrak{J}, \vee, \wedge, \sim)$ is called a concept Boolean algebra.

Now, we need to know how do describe basic concepts in factor space.

Definition 6 A concept $\alpha = (E, A)$ is called a *basic concept* if the intension A can be stated by a sentence with the disjunction formal form (FDCF):($f_1(u)$ is A_1)AND…AND($f_n(u)$ is A_n). Where A_j is a subset of $X(f_j)(j = 1, \ldots, n)$.

It is obvious that the definition can be stated as follows equivalently:

Definition 3 ′ [14] A concept $\alpha = (E, A) \in \mathfrak{J}$ is called a *basic concept* if

1. $A = A_1 \times \cdots \times A_n \subseteq G$ is a (super)rectangle in $X(f_1) \times \cdots \times X(f_n)$;
2. The rectangle can not be extended within G: There is no $B = B_1 \times \cdots \times B_n \subseteq G$ such that $B \supseteq A$ and $B \neq A$.

It is very easy to prove the following propositions:

Proposition 2 *Given a factor space $\psi = (\{X(f)\}(f \in F); U)$, let B be the set of all basic concepts. $(B, \wedge)$ is a concept semi-lattice.*

What is the objective of concept semi-lattice searching? Is it best that pick out all sub semi-lattices from $(B, \wedge)$ No, the number of sub semi-lattices is too big for use, especially in big databases. Yes, what we want is dividing concepts as fine as possible, but the fewer middle concepts the better. According to this objective, factor space theory presents the principle of maximum discernibility for concept forming [11], and the principle of maximum degree of determination for decision making [13].

So far, we can see that the concept semi-lattice has been described by factor space clearly. Is the involution kept in concept definition in factor space? Yes, it is!

Now, we are going to discuss the involution of intension and extension in factor space.

Definition 7 We call a concept $\alpha = (E, A) \in \mathfrak{J}$ satisfying involution if $F^{-1}[F(E)] = E$.

Proposition 3 *Any basic concept $\alpha = (E, A) \in \mathfrak{I}$ satisfies involution.*

Proof Firstly, we need to get the general form of extension of basic concepts. According to the Definition 5, $\alpha = (E, A) = (F^{-1}(A), A)$. Since it is a basic concept, $A = A_1 \times \cdots \times A_n \subseteq G$. Suppose that $A_j = a_{j1} + \cdots + a_{jn(j)}, a_{jk}$ are quality-atoms in $X(f_j)$ for $j = 1, \ldots, n$ and + stands for union operation on disjoint subsets: $a_{jk} \cap a_{jl} = \emptyset (k \neq l)$. We have that

$$F^{-1}(A) = F^{-1}(A_1 \times \cdots \times A_n) = \cap f^{-1}(A_j) \mid j = 1, \ldots, n$$
$$= \cap\{\cup\{[a_{jk}] \mid k = 1, \ldots, n(j)\} \mid j = 1, \ldots, n\}$$

This is the general form of extension of basic concepts. Given $E = \cap\{\cup\{[a_{jk}] \mid k = 1, \ldots, n(j)\} \mid j = 1, \ldots, n\}$, we have that

$$F(E) = \wedge\{\wedge\{[a_{jk}] \mid k = 1, \ldots, n(j)\} \mid j = 1, \ldots, n\} = \mathbf{a}$$

$$F^{-1}(F(E)) = F^{-1}(\mathbf{a}) = F^{-1}(\wedge\{\wedge\{a_{jk} \mid k = 1, \ldots, n(j)\} \mid j = 1, \ldots, n\})$$
$$= \cap\{\cup\{[a_{jk}] \mid k = 1, \ldots, n(j)\} \mid j = 1, \ldots, n\} = E$$

It means that $\alpha = (E, A)$ satisfies involution.

The definition of involution has a little changing from the original definition given by Wille. We just request that $F^{-1}[F(E)] = E$ here but do not require that $F[F^{-1}(A)] = A$. The reason is: From upper-seat concept to lower-seat concept, we often omit to state those attributes inherited from the upper-seat but only new attributes the lower-seat concept creates.

4 Conclusion

The context concept analysis has important applications in intelligent artificial. This paper contributes some points to improve the theory under the theory of factor space. Concept analysis can be done perfectly provided we put the problem into the frameworks of factor space. The main topic in context analysis is the forming of concept and the concept semi-lattice. The tableaus of formal background are too complex since the tableaus' design is attribute-oriented. Factor space uses simple tableau by factor-oriented. The practical algorithms on context analysis are complex, some of them are N-hard problem, while factor space has simple and fast algorithms to realize concept-forming. (See papers [11, 12]).

References

1. Ganter, B., Wille, R.: Formal Concept Analysis. Springer, Berlin (1999)
2. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival, I. (ed.) Ordered Set, pp. 445–470. Reidel, Dordrecht-Boston (1982)
3. Wang, P.Z., Sugeno, M.: The factors field and background structure for fuzzy subsets. Fuzzy Math. **2**, 45–54 (1982)
4. Wang, P.Z., Li, H.X.: A Mathematical Theory on Knowledge Representation. Tianjin Scientific and Technical Press, Tianjin (1994)
5. Yuan, X., Wang, P.Z., Lee, F.S.: Factor space and its algebraic representation theory. J. Math. Anal. Appl. **17**(1), 256–276 (1992)
6. Wang, P.Z.: A factor spaces approach to knowledge representation. Fuzzy Sets Syst. **36**(1), 113–124 (1990)
7. Foong, S.B., Wang, P.Z.: A factor space approach to concept representation. In: Proceedings of Fuzzy Logic in Artificial Intelligence, pp. 97–113 (1993)
8. Li, H.X., Wang, P.Z., Yen, V.C.: Factor spaces theory and its applications to fuzzy information processing (I). The basics of factor spaces. Fuzzy Sets Syst. **95**(2), 147–160 (1998)
9. Li, H.X., Chen, C.L.P., Yen, V.C. et al.: Factor spaces theory and its applications to fuzzy information processing: Two kinds of factor space canes. Comput. Math. Appl. **40**(6–7), 835–843 (2000)
10. Zou, K.: Basic Algebra and Category Theory. Dalian Maritime University Press (2002)
11. Wang, P.Z.: Factor spaces and factor data-bases. J. Liaoning Tech. Univ.: Nat. Sci. **32**(10), 1–8 (2013)
12. Wang, P.Z., Guo, S., Bao, Y.K., Liu, H.T.: Factorial analysis in factor space. J. Liaoning Tech. Univ.: Nat. Sci. **32**(10), 1–8 (2013)
13. Liu, H., Guo, S.: The reasoning model for factorial analysis. J. Liaoning Eng. Tech. Univ. (2015)
14. Wang, P.Z., Liu, Z.L., Shi, Y., Guo, S.C.: Factor space, the theoretical base of data science. Ann. Data Sci. **1**(1) (2014)

An Improved Genetic Algorithm for Multiple Sequence Alignment

Man-zhi Li, Hai-xia Long and Cheng-yi Zhang

Abstract Multiple sequence alignment (MSA) is one of the most essential tools in bioinformatics. Genetic algorithm is used to simulate biological multiple sequence alignment problem, the initial population and crossover is the most critical part of the genetic algorithm. In this paper, we construct three initial populations and a simple horizontal crossover with respect to the vertical crossover. The experimental results showed that the initial population adding an appropriate proportion MAFFT excellent seed can optimize the population, horizontal crossover can reduce the computing time and the computational complexity. Combination of those two methods can improve the computational efficiency of multiple sequence alignment.

Keywords Multiple sequence alignment · Genetic algorithm · MAFFT seeds · Horizontal crossover · Computational efficiency

1 Introduction

Multiple sequence alignment problem (MSA) is an unresolved problem in bioinformatics. It belongs to a NP-complete combinatorial optimization problem, since finding the exact polynomial complexity algorithm is not possible [1]. Therefore, finding approximate solution of the multiple sequence alignment problem have practical significance. Genetic algorithm (GA) have been successfully applied to solve MSA problem [2–6], it borrows the views biogenetics, and achieves to improve the adaptability of each individual by selection, crossover, mutation and other operators.

Genetic algorithm uses an initial population and tries to improve the population using mutation and crossover operators. The initial population is a key part of GA. It is usually created by randomly inserted gaps in the sequence. It has no more than

M. Li · H. Long (✉) · C. Zhang
School of Mathematics and Statistics, Hainan Normal University,
Haikou 571158, Hainan, China
e-mail: haixia_long@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_41

biological significance and makes the algorithm computationally expensive. The quality of initial population directly affects the convergence speed, the population with high fitness value can quickly converge to the optimal solution [7, 8]. MSA have many common online testing tools such as MAFFT, CLUSTALW, MUSCLE, etc., their results, although not the same, still can be used as a excellent reference. With different parameters input, MAFFT alignment results are different. If these MAFFT results added to the initial population, the quality of the initial population will be optimized.

Genetic operators is another key part of GA. Notredame et al. [3] set up crossover, adding gap, mutation and other 22 genetic operators; Naznin et al. [6] also applied single point crossover and multiple point crossover; Fan et al. [9] proposed a smart operator in GA. Some genetic operator structure are more complex, but Thomsen and Boomsma [10] verified a simple or complex crossover for the alignment results were not significantly different. Goondro and Kinghorn [7] considered that over complex operator is unnecessary designed, GA can achieve MSA basic requirements, so single point crossover becomes the most popular crossover. According to the crossover path, single point crossover belongs to the overall vertical crossover, MSA require that residue must remain the same order after crossover, so vertical crossover have difficulty of programming. If we consider single point horizontal crossover, you can reduce the difficulty of programming.

In this paper, GA is applied for solving MSA. Initial population and operators are important elements of GA, their effect on the performance of the algorithm and quality of result are discussed. Online alignment tool MAFFT results are added in the initial population, which can optimize the quality of the population, propose single point horizontal crossover and compared with single point vertical crossover. The experimental results showed that by adding an appropriate proportion MAFFT excellent seed in the initial population, the population can be optimized, horizontal crossover can reduce the computing time and the computational complexity. Combination of this two methods can improve the computational efficiency of multiple sequence alignment.

2 Multiple Sequence Alignment Problems and Mathematical Description

Given a family $S = (s_1, s_2, \ldots, s_k)$ of k sequences of various lengths n_1 to n_k and a finite alphabet set $\sum$. Each sequence element represents a character from a given alphabet (for DNA sequences, the alphabet consists of four characters of nucleotide {A, T, C, G}. For protein sequences, the alphabet consists of 20 characters of the amino acids). An alignment of the sequence family S' is a $k \times N$ matrix. $S' = (s'_{ij})$, where $s'_{ij} \in \sum \cup \{‘-’\}$, and ‘−’ indicates a gap in the alignment as the result of an insertion or deletion of an nucleotide base. A new sequence set S' is called an alignment of S if it satisfies the following properties:

(i) The strings in S' have the same length.
(ii) Omitting gaps, sequence S_i' is identical to sequence S_i.
(iii) S' has no column which only contains gaps.

3 Design and Analysis of Algorithms

3.1 Coding Strategy

An alignment can be interpreted as an array with n rows and m columns, and each row contains one sequence with extra gap. We use two-dimensional chromosome coding. Table 1 gives an example with six sequences.

3.2 Population Initialization

3.2.1 Random Initialization

Several gaps are random inserted in the sequences, the sequences length which have inserted gap is not more than 1.2 times of the longest sequences length without gaps. Following these steps:

- Determining the length of the sequence inserted gap $L = [len_{\max} \times 1.2]$, where $len_{\max}$ is the maximum length of each sequence, that is the length of the longest sequence.
- The number of gaps inserted in the ith sequence is $L - len_i$. Create random repeatable position, and insert the gap into the appropriate position.
- According to the original series, the characters are copied in the order of the non-gap appropriate location.
- If there are all gaps in column, delete the gap column.

Table 1 Example of six sequences alignment

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I	Y	Y	D	G	G	A	V	–	E	A	L	C	A	M
II	Y	D	D	–	G	A	L	V	E	A	L	C	A	M
III	F	D	E	G	G	–	L	V	Q	A	–	C	F	
IV	F	Y	E	G	G	I	V	V	Q	A	V	–	–	M
V	Y	D	D	G	G	I	L	V	–	–	L	C	A	M
VI	Y	Y	E	–	G	A	–	V	Q	A	A	C	E	M

3.2.2 Add MAFFT Seeds in Initial Population

When we change different parameters which include gap open penalty (gop) and gap extension penalty (gep) in the online alignment tool MAFFT, we can get different results. Assume the size of the initial population is 50, we input 50 group different parameters gop: gep, then get 50 MAFFT alignment results as the good seeds.

If these 50 excellent seeds are all the individuals in original random initial population, this population is called 50 MAFFT seeds.

If 20 excellent seeds are randomly selected from these 50 MAFFT seeds, and randomly replaced 20 individuals in original random initial population, this population is called 20 MAFFT seeds.

If the original random initial population is not inserted in any excellent seeds, it is called 0 MAFFT seeds.

3.3 *Fitness Function*

The Sum of Pairs (SP) is commonly used as a fitness measure for MSA, SP is as follows [11]:

$$SP = \sum_{h=1}^{L} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} Cost(S_i, S_j)$$

Here, SP is the cost of MSA. L is the length of the alignment, k is the number of sequences, $Cost(S_i, S_j)$ is the alignment score between the two aligned sequences S_i and S_j. When $S_i \neq$ '−' and $S_j \neq$ '−', $Cost(S_i, S_j)$ is determined by substitution matrix which common PAM or BLOSUM series. When S_i = '−' and S_j = '−', $Cost(S_i, S_j) = 0$. When S_i = '−' and $S_j \neq$ '−' or $S_i \neq$ '−' and S_j = '−', we use a model with affine gap penalties as shown in

$$penalty = gop + n \cdot gep$$

Here, penalty is the gap penalty, gop is the cost of opening a gap, gep is the cost of extending the gap by one, n is the length of the gap.

3.4 *Selection Operator*

Algorithm uses two types of selection operators: elitist method and roulette wheel method, individual's fitness values are sorted, the top 10 % of the individuals are retained until the next generation, and the remaining 90 % individuals are applied by roulette wheel selection method to select the parent.

3.5 Crossover Operator

The pairs are randomly selected with crossover probability in the population, according to the crossover path, it is divided into single point vertical crossover and single point horizontal crossover.

Single point vertical crossover:
A crossing point is random set in the parent1 individual, and find the corresponding crossing point in the parent2 individual, exchange front or later of the crossing point parts, delete the gap column, generate two new child individuals. It is implemented as shown in Fig. 1.

Single point horizontal crossover:
Create a random which is less than the individual rows number, exchange this row in two parent individuals, delete the gap column, generate two new child individuals. It is implemented as shown in Fig. 2.

3.6 Mutation Operator

Using one bit mutation method. The individual is randomly selected with mutation probability in the population, and select a randomly location in a strip sequence, if this position is a gap, then randomly choose a non-gap location, the gap is moved to

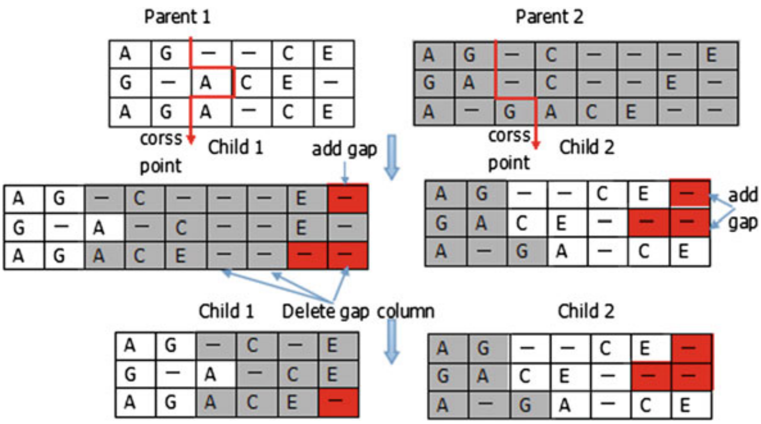


Fig. 1 Single point vertical crossover

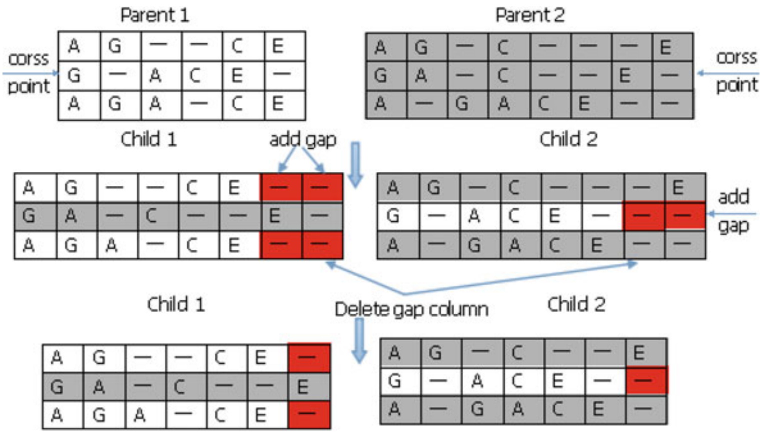


Fig. 2 Single point horizontal crossover

this position, the other characters to the left or right, delete the gap column, generate new child individual.

4 Experimental Results and Analysis

4.1 *Experimental Parameters*

See (Table 2).

4.2 *The Evaluation Criteria of Alignment Results*

The BAliBASE 2.01 (Benchmark Alignment dataBASE) was used to perform experiments. This database has a total of 141 benchmarks [12]. To assess the performance of the programs in this study, we calculate scores which estimate the

Table 2 Experimental parameters table

Parameter	Dataset10	Fitness function	Sum-of-pairs
Population size	50	Score matrix	BLOSUM45
Generations	3000	Crossover operator	One point
Selection strategy	Roulette wheel	Crossover probability	0.6
Elitist rate	0.1	Mutation operator	One bit
Programming language	MATLAB	Mutation probability	0.2
Coding	Two-dimensional	gop:gep	3:0.15

quality of an alignment compared to the BAliBASE reference. The sum-of-pairs score (SPS) is calculated such that the score increases with the number of sequences correctly aligned. It is used to determine the extent to which the programs succeed in aligning some, if not all, of the sequences in an alignment [13].

Suppose a test alignment of N sequences consisting of M columns. We can designate the ith column in the alignment by $c_{i1}, c_{i2}, \ldots, c_{iN}$. For each pair of residues c_{ij} and c_{ik} define P_{ijk}:

$$P_{ijk} = \begin{cases} 0 & c_{ij} = c_{ik} \\ 1 & c_{ij} \neq c_{ik} \end{cases}$$

S_i is defined as:

$$S_i = \sum_{j=1, j \neq k}^{N} \sum_{k=1}^{N} P_{ijk}$$

The SPS for the alignment is:

$$SPS = \sum_{i=1}^{M} S_i / \sum_{i=1}^{M_{ref}} S_{ref}$$

where M_{ref} is the number of columns in the reference alignment and S_{ref} is the score S_i for the ith column in the reference alignment. The SPS indicates the ratio of pairs correctly aligned.

4.3 Result and Analysis

10 datasets are randomly selected from BAliBASE2.01, the SPS is the quality evaluation standards of alignment results. Due to the randomness of the genetic algorithm, the same experiments parameters may give different results, so this example uses the same method and runs 10 times. We calculate the SPS value, their average value and the maximum value as the final result, 0M is 0 MAFFT seed, and the same meaning as 20M, 50M (Figs. 3 and 4, Tables 3 and 4).

For all the tests, we can see that:

- The computing time of horizontal crossover is shorter than that of vertical crossover, and the SPS of horizontal crossover is better than vertical crossover. So the computational efficiency of horizontal crossover is better than vertical crossover.
- When sequences length <120 bp, the SPS of 0 MAFFT seed is nearly with 20 MAFFT and 50 MAFFT, so the best choice is 0 MAFFT.

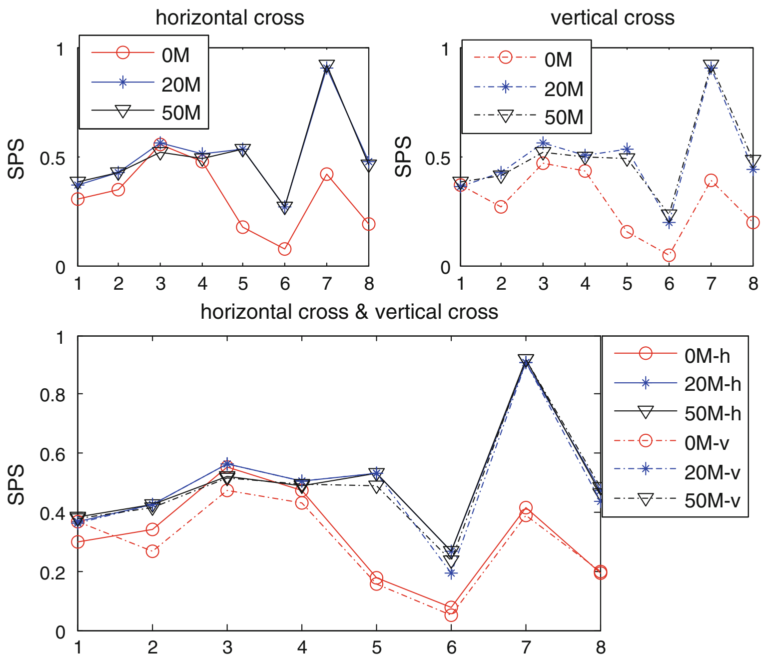


Fig. 3 Average SPS of three initial populations and two crossovers

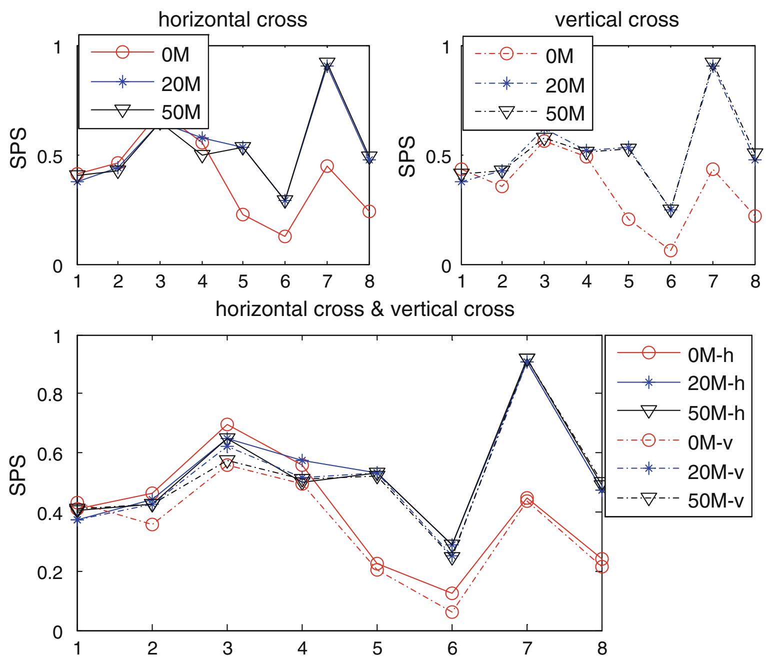


Fig. 4 Best SPS of three initial populations and two crossovers

Table 3 Average SPS of proposed method on BAliBASE2.01 dataset (black is max SPS in this dataset)

Name of dataset	Length (bp)	SPS (horizontal crossover)			SPS (vertical crossover)		
		0M	20M	50M	0M	20M	50M
451c_ref1	87	0.3005	0.3668	0.3851	0.3661	0.3616	0.3801
1aab_ref1	79	0.3436	0.4265	0.4246	0.2682	0.4246	0.4139
1aho_ref1	65	0.5549	0.5613	0.5212	0.4714	0.5632	0.5184
1pfc_ref1	116	0.4742	0.5076	0.4895	0.4316	0.5065	0.4956
2cba_ref1	259	0.1781	0.5325	0.5325	0.1549	0.5325	0.492
2pia_ref1	287	0.0785	0.2687	0.2683	0.0486	0.1939	0.2337
5ptp_ref1	245	0.4162	0.9066	0.9165	0.3897	0.9066	0.9165
kinase_ref1	276	0.1922	0.4755	0.4622	0.1959	0.4361	0.4823
Time (s)		7800			10000		

Table 4 Best SPS of proposed method on BAliBASE2.01 dataset (black is max SPS in this dataset)

Name of dataset	Length (bp)	SPS (horizontal crossover)			SPS (vertical crossover)		
		0M	20M	50M	0M	20M	50M
451c_ref1	87	0.4096	0.3740	0.4041	0.4329	0.3753	0.4082
1aab_ref1	79	0.4623	0.4397	0.4246	0.3568	0.4246	0.4246
1aho_ref1	65	0.6957	0.6462	0.6479	0.5590	0.6205	0.5726
1pfc_ref1	116	0.5559	0.5745	0.4981	0.4926	0.5158	0.5084
2cba_ref1	259	0.2223	0.5325	0.5325	0.2015	0.5325	0.5216
2pia_ref1	287	0.1236	0.2875	0.2875	0.0614	0.2487	0.2436
5ptp_ref1	245	0.4495	0.9066	0.9165	0.4341	0.9066	0.9165
kinase_ref1	276	0.2386	0.4755	0.4892	0.2153	0.4755	0.5016
Time (s)		7800			10000		

- When sequences length >120 bp, the SPS of 0 MAFFT seed is greatly less than 20 MAFFT and 50 MAFFT, 20 MAFFT and 50 MAFFT are appropriate. However, the SPS of 50 MAFFT is nearly with 20 MAFFT, but this higher complexity is offset by better quality results, and it cause the algorithm to entrap in local optima, so the best choice is 20 MAFFT.
- For random algorithm such as GA, although the average data may reflect the performance of the algorithm, the final result is still subject to the maximum, and the trend of average and maximum is similar. Thus, we only compute the maximum in GA henceforth.

As a conclusion, we can state a good configuration for future tests, short sequences alignment can choice 0 MAFFT and horizontal crossover, long sequences alignment can choice 20 MAFFT and horizontal crossover.

5 Conclusion

In this paper, the adopted methodologies use genetic algorithm to analyze three initial populations and two crossovers for evaluating MSA that improve traditional alignment quality. The aim is to find the combination of these metrics that achieves the best approximation to the biological quality of the MSA.

We proposed a new method for population initialization by randomly adding an appropriate proportion MAFFT excellent seed to the initial population. The advantage of this method is that it does not have the problem of getting entrapped in the local optima of seed. This final answer is highly dependent on its initial population. In fact a good initial population leads to a good result. The new initial population showed better efficiency on medium sequences. We proposed a new single point horizontal crossover which has low computational complexity and shorter computing time than traditional vertical crossover.

The experimental results showed that adding an appropriate proportion MAFFT excellent seed to the initial population can optimize the population, and horizontal crossover can reduce the computing time and the computational complexity. Combination of those two methods can improve the computational efficiency of multiple sequence alignment.

For future work we can use our new combination to solve the problem of large scale sequences alignment.

Acknowledgments This work was supported by the National Natural Science Foundation 71361008, Hainan Province key projects ZDXM20130080, Hainan Province Natural Science Fund 20151003.

References

1. Wang, L., Jiang, T.: On the complexity of multiple sequence alignment [J]. J. Comput. Biol. **1** (2), 337–348 (1994)
2. Notredame, C., Higgins, D.G.: SAGA: sequence alignment by genetic algorithm [J]. Nucleic Acids Res. **24**(8), 1515–1524 (1996)
3. Notredame, C., O'Brien, E.A., Higgins, D.G.: RAGA, RNA sequence alignment by genetic algorithm [J]. Nucleic Acids Res. **25**(22), 4570–4580 (1997)
4. Kaya, M., Sarhan, A., Alhajj, R.: Multiple sequence alignment with affine gap by using multi-objective genetic algorithm [J]. Comput. Methods Prog. Biomed. **114**(1), 38–49 (2014)
5. Ben Othman, M.T., Abdel-Azim, G.: Genetic algorithms with permutation coding for multiple sequence alignment [J]. Recent Pat. DNA Gene Sequences **7**(2), 105–114 (2012)
6. Naznin, F., Sarker, R., Essam, D.: Progressive alignment method using genetic algorithm for multiple sequence alignment [J]. IEEE Trans. Evol. Comput. **16**(5), 615–631 (2012)
7. Gondro, C., Kinghorn, B.P.: A simple genetic algorithm for multiple sequence alignment [J]. Genet. Mol. Res. **6**(4), 964–982 (2007)
8. Narimani, Z., Beigy, H., Abolhassani, H.: A new genetic algorithm for multiple sequence alignment [J]. Int. J. Comput. Intell. Appl. (4), 1–18 (2013)
9. Fan, H., Wu, R., Liao, B., et al.: An improved genetic algorithm for multiple sequence alignment [J]. J. Comput. Theory Nanosci. **9**(10), 1558–1564 (2012)

10. Thomsen, R., Boomsma, W.: Multiple sequence alignment using SAGA: investigating the effects of operator scheduling, population seeding, and crossover operators [M]. In: Applications of Evolutionary Computing, pp. 113–122. Heidelberg, Springer (2004)
11. Gusfield, D.: Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology [M]. Cambridge university press (1997)
12. Bahr, A., Thompson, J.D., Thierry, J., et al.: BAliBASE (Benchmark Alignment dataBASE): enhancements for repeats, transmembrane sequences and circular permutations [J]. Nucleic Acids Res. **29**(1), 323–326 (2001)
13. Thompson, J D, Plewniak, F., Poch, O.: BAliBASE: a benchmark alignment database for the evaluation of multiple alignment programs [J]. Bioinformatics **15**(1), 87–88(2) (1999)

The Composition and Resolution of Fuzzy Linguistics in Factors Space

Jing Zou and Zewen Jia

Abstract In order to apply Mr. Wang's factor space theory to application easily, the author puts forward the fuzziness of natural language based on the factor space theory. I further elaborate the research on the composition and resolution of fuzzy linguistics and point out that the ordinary composition chart has been replaced by new generalized one. This is to convey the composition and resolution of fuzzy linguistic more properly so that the research of fuzzy linguistics can be advanced.

Keywords Factor space · Fuzzy linguistics · Natural language

1 Introduction

The factors space theory has been developed greatly since its foundation by Mr. Wang Peizhuang, whose latest paper described and discussed the concrete concept of factors space and factors pool. The author is good at studying the problems in the linguistics domain. In the past year, the research of linguistics was carried out under this leading idea: Language should be fairly precise. But in many occasions, the content expressed in nature language isn't precise as people have expected. It is difficult to deal with language in an accurate way in many circumstance. Theoretical linguists often meet with such problem during their research: many things which

The article is National Natural Science Foundation of China (NSFC)'s phased objectives in "Mechanism Study of Neural Network Fuzzy Inversion and Crime Dark Space" (Project No. 61272170).

J. Zou (✉) · Z. Jia
College of English and Construction, Dalian University, Dalian, Liaoning 116622, China
e-mail: zoujing1778@126.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_42

can be precisely distinguished in theory are difficult to prove accurately once they meet language example. In the long-term analysis of practical linguistics study, such problems were encountered that can't be solved by traditional method. In 1965, the expert in the cybernetics at the University of California, Berkeley, L.A. Zadeh founded fuzzy mathematics in his pioneering work "Fuzzy Sets". In recent years, fuzzy mathematics has been developed quickly and applied in all kinds of fields. It has also been paid great attention in interdisciplinary studies. There is no doubt that Fuzzy Linguistics came into being under this background.

The fuzziness of natural language is common occurrence in English. It is expressed in many parts of speech such as noun, verb, adj., adv., etc. This fuzzy linguistics has some features: firstly, it is a multi-factor decision system; secondly, the expression of elements is a threshold, namely the factor is a scope; thirdly, the essential data derives from many linguists' experience, making their evaluation base different as a result of different linguistics context; fourthly, the essential data is the personal description of a certain object, making it ambiguous; fifthly, how to ascertain the confidence degree of these description and whether these descriptions can proof mutually.

In order to solve the above problems, the author realized that sometimes dealing with natural language in a fuzzy way is more effective than in a precise way after the birth of factor space proposed in Mr. Wang's literature [1]. Language presents fuzziness in many aspects. This sort of fuzziness is not a kind of shortcoming of language as people have thought, but a basic attribute in the complicated system of language. Therefore it can not be avoided.

2 Prerequisite Knowledge

Factor is the element of analyzing object property and causal association. Factor space is the coordinate space named by factors, which is the universal mathematical frame of object description. It is the basic mathematical theory of artificial intelligence, intelligent data science in particular.

Factors are defined mathematically as a kind of mapping. It maps one object into a property value, and such mapping is called qualitative mapping; meanwhile, mapped it into a condition, such mapping is called quantitative mapping. For example, stature is a mapping, and it can map Zhang San into condition "very high" as well as into a quantitative condition "1.8 m", shown as Fig. 1. Everything has two prescriptions of quality and quantity simultaneously. From quantitative change to qualitative change, quantity determines quality.

Based on this philosophy, supposed these two mappings coexist. The quantitative mapping of factor f forms a coordinate axis X_f, the attribute value obtained from quantitative mapping of factor f is an ordinary subset or fuzzy subset of the formation of fuzzy subset has been solved by fuzzy set theory.

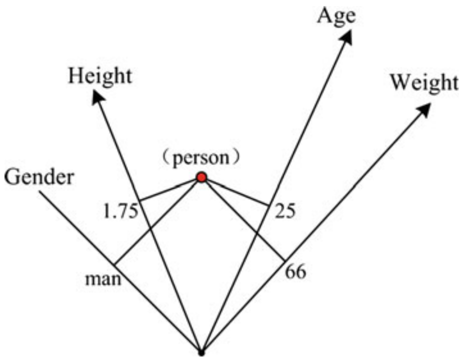


Fig. 1 A Quantitative Condition

Factors are analysis, which is abstract. Back to specific is the analysis combined process of cognition. As shown in the figure, join up these factor axes and get the coordinate frame named by factor, which is factors space. Anything can be considered as a point in the factor space. The mathematical definition is that a family of sets whose index set is factor *F*. Hereby *F* is a set of various factors and it is a Boolean algebra as there exists Boolean calculation among factors. Therefore, factors space can be defined as a family of sets whose index set is Boolean algebra.

As a kind of new database, factor base is the data implementation of factors space theory and its data treatment adopts a series of tabular form.

Applying factors space on the study and treatment of factor base not only can conduct the property division and concept lattice extraction from its conceptual analysis sheet, but also has more advanced theory and application. The main research phenotypes include: (a) factor analysis sheet; (b) multi-objective decision table for the optimization of objective factors; (c) safety warning table for the assessment of safety factors; (d) stability adjustment table for the adjustment of governing factors; (e) the establishment of factors dictionary and factors header base; (f) the fuzziness of natural language.

3 The Fuzziness of Natural Language

The fuzziness of natural language is common occurrence in English. It is expressed in many parts of speech such as noun, verb, adj., adv., etc. Many scientists have done experiments on the fuzziness of noun. Among these it's the famous experiment of [2, 183] on the English word of bird. He asked people who took part in his experiment to classify the identical degree of the mentioned animals with the meaning of "bird". The result proved that, when judging whether an animal was a bird, sometimes it is difficult to make a simple answer with a single "yes" or "no". The extension of "bird" has unity and coherence; Swallows are indisputable birds, chickens and other poultries are secondary. While penguins and ostriches are more non-affirmative. As for bats there lies much more doubtful points. To a rigid extent, bats are not birds. But because they have some characters of birds, for example, they have wings and they can hardly regarded as birds. From this, there is unity and coherence in the meaning of birds. In fact, there are similar phenomena in many other nouns. The limits of the range of the words meaning are not entirely different. Like birds, diseases, children and toys, although the intension of their meaning can be smoothly defined in abstract terms, their extension is difficult to be clearly classified. This point produces the fuzziness of the language. In addition, as all we know, baldhead paradox applies the fuzziness of the language, too. A person with no hair is a baldhead. A person with a hair is also a baldhead, because he has one more hair than the former. A person with two hairs should be a baldhead because he also has one more hair than the former. And so on and so forth, we can conclude that if a person with n hairs is a baldhead, a person with n + 1 hairs is also a baldhead. Therefore all people are baldheads. This is an obvious paradox. The reason lies in the fuzziness of the word "baldhead", that is to say, there is no clear limit between a baldhead and a non-baldhead, and we can not classify it using the specific number of hairs.

4 The Composition and Resolution of the Fuzzy Linguistics

In linguistics, to define some words with digits so as to make them fixed quantitatively is a necessary job to make electronic computer understand natural language.

Word is the smallest unit in natural language. We call the basic indivisible word atom word such as person, cat, horse, beautiful, ugly, quick, slow etc. The set of words is considered as universe of discourse "U" Semantics is expressed through N, the corresponding relation of T to U. Generally speaking, N is a fuzzy relation. As for any fixed a, a ϵ T, it follows:

$$N(a, u) = \mu A(u)$$

It is a fuzzy subset of "U", writing it down as A. Especially when A = A, that is, it is an ordinary set, the word "a" is called "clear", or it is called "fuzzy".

Phrases are composed of words. From the angle of mathematical logic, phrases are composed of words and logic operational symbols "∧", "∨", "¬", "∧", "∨", "¬" are corresponded with "∩", "∪" in set theory and complementary operation "¬". Thus, words can be made up into phrases or phrases can be divided into words. If we bracket words, it follows:

(1) Asia-Africa-Latin America = [Asia] ∪ [Africa] ∪ [Latin America]
(2) White bird = [White] ∪ [bird]
(3) Non-animal = ¬[animal]

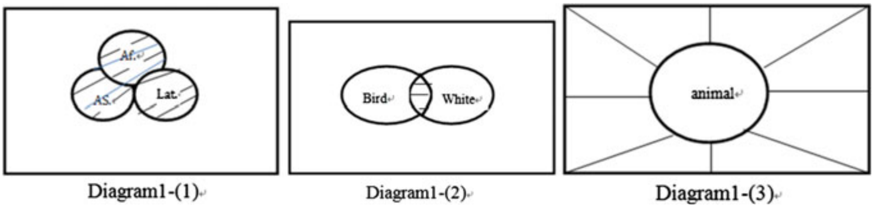


Fig. 2 Venn Diagram

These words can be shown with Venn Diagram (Fig. 2).

Attention: Because we apply fuzzy linguistics, the above showing way of Venn Diagram is not very accurate. For example, "white" has not accurate limits, while "animal" has not a clear range. So we can not show them with entirely different general sets. We must apply extended Venn Diagram of fuzzy limits. This is the composition of fuzzy language. For example, in the Fig. 3, suppose "A" stands for the fuzzy subset of "birds", "B" stands for white fuzzy subset. It follows:

In the end, from the composition and resolution of linguistics variable, we conclude the definition of fuzzy linguistics variable:

Language variable can be defined as the following quinary group (X, T(X).U, G, M)

(1) X is the name of language variable, such as young or old of age and number etc.
(2) T[X] is the set of language true value, for example:

$$\begin{aligned} \mathrm{T(X)} &= \mathrm{T}(age) \\ &= very\ young + young + middle - age + relatively\ old + very\ old. \end{aligned}$$

(3) U is universe of discourse, when "X" is "age", we consider U as [0, 100], the unit as "year of age."
(4) G is the grammar rule, achieving the membership function of phrases according to atom words.
(5) M is semantic rule, achieving the membership function of fuzzy subset X according to semantic rules.

Taking age as an example, this quinary group can be seen as followed (Fig. 4):

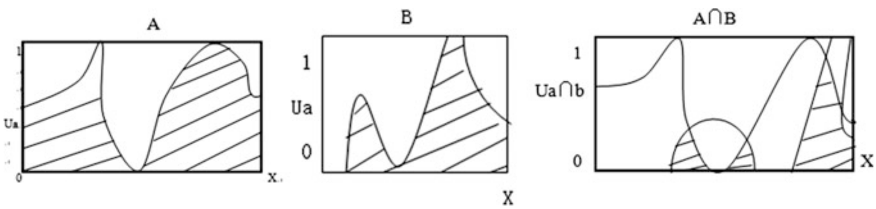


Fig. 3 Extended Venn Diagram

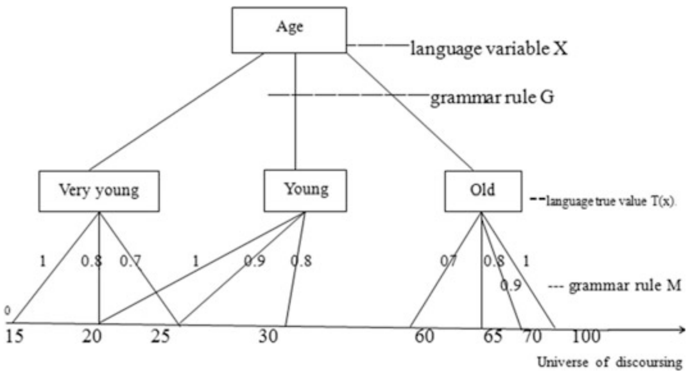


Fig. 4 Quinary group

5 The Conclusion

The author discussed the attribute expression method of factors space object that proposed by Mr. Wang, which points out the fuzziness of natural language. I do further research on the composition and resolution of fuzzy linguistics. Sometimes dealing with natural language in a fuzzy way is more efficient than in a precise way. The effective communicative effects can only be realized by appropriate uses of linguistic hedges.

This sort of fuzziness isn't a kind of defect of language as people have considered, but a basic attribute in the complicated system of language.

References

1. Wang, P.: Factor spaces and factor databases. J. Liaoning Tech. Univ. (Natural Science) **32** (10), 1–8 (2013)
2. Lakoff, G.: Hedges: A study in meaning criteria and the logic of Fuzzy. Concepts Proc. Chicago Linguist. Soc. **8**, 183–228 (1972)
3. Wang, P.: Fuzzy Set Theory and Its Application. Tianjin Science and Technology Publishing House (1983)
4. He, Z.: Fuzzy Mathematics and Its Application. Tianjin Science and Technology Publishing House (1983)
5. Wang, Y.: Cognitive Linguistics. Shanghai Foreign Education Publishing House (2007)
6. Wang, Y.: Cognitive and discourse analysis. Foreign Educ. Res. (2003)
7. Cui, T., Ma, Y.: Research on multi-dimensional space fault tree construction and application. China Saf. Sci. J. **23**(4), 32–37 (2013)
8. Yi, Q., Lin, Y., Wu, G.: Research of fusing image classification into object localization. Appl. Res. Comput. (2013). http://www.cnki.net/kcms/detail/51.1196.TP.20130730.0909.003.html
9. Death Knell: Technology as a Passport to Personalized Education. http://www.nytimes.com/2011/12/06/science/daphne-kOller-teChllOlOgy-as-a-passport-t0-persOllahZed-ednCat1011.html?pagewanted=1}r=2,2013-12-15

10. Open Education: Online education-Introducing the microlecture format [EB/OL] (2009). http://www.openeducation.net/2009/03/08/online-education-introducing-the-microlecture-format/
11. Zadch, L.A.: Quantitative fuzzy semantics. Inf. Sci. **3**(2) (2014)

Research on SCADA Security Defense Model Based on Petri Nets

Fan Yang, Xie-dong Cao, Meng Zhao and Wei-wei Zhang

Abstract Through research and analysis of difference between the security risk of SCADA system of oil and gas gathering and transportation and general information system, combined with factors neural network theory, the project team proposed the security and defense model of SCADA system based on FNN. This paper focuses on formal description of model by Petri nets, using fuzzy Petri nets to complete the description of internal reasoning of executive neurons.

Keywords SCADA · Petri nets · Factor neural network · Formal description

1 Introduction

Petri nets is a kind of system modeling method represented by mesh graphical, which has a strict mathematical definition of graphical modeling tool, and has a strong function distribution and the functional description ability, is a powerful tool system modeling and performance analysis for dynamic system of asynchronous, discrete and concurrent events. It has been widely used in many fields. Petri nets modeling has two methods: direct modeling and indirect modeling, this paper adopt the second method, model built by factor neural network theory is transformed into Petri nets model, the Petri nets is used as a kind of formal modeling method. The reason is that the original model is too abstract, the model can be described by graphical description ability of Petri nets more clearly.

F. Yang (✉) · X. Cao · M. Zhao · W. Zhang
School of Electrical Engineering and Information,
Southwest Petroleum University, Chengdu 610500, China
e-mail: yfsy2010@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_43

2 FNN Model of SCADA Security Defense System

2.1 Model Requirement Analysis

The project this paper belongs to intends to establish the SCADA security and defense model by using factor neural network as knowledge transfer, requiring to record malicious programs and operation of each host in different space and different time, analyzing and deciding the shutdown of instrument and remote terminal unit of the current host and the protective measures against the host, which includes a storage and release of factors knowledge, management, security event statistics, event tracking analysis, safety evaluation, and early warning systems and so on.

2.2 Behavioral Factors of Malicious Programs

Achievement of function of malicious programs always have to rely on a series of program behavior sequence, the relationship between the behavior of the program and the program is the relationship between the individual 1 and the general, the realization of specific behavior relates to call various application programming interfaces(API) provided by the operating system.

The program behavior can be monitored by HOOK API technology, and then the program is decomposed indirectly [1]. The corresponding program is represented as a behavior set: Program = {Oper1, Oper2, … Operi}, ∀Operi ∈ Program (i = 1, 2, …, n), Oper (Operation) represents specific behavior.

Because there is no a clear one to one relationship between the malicious programs and its behaviors, which resulted ambiguity between program behaviors and the programs. Fuzzy reasoning is a kind of effective method of uncertain reasoning, the API function of the key suspicious behavior is extracted as the factor to construct the fuzzy inference factor neuron [2].

2.3 Factor Neural Network Model of SCADA System

According to model requirement analysis and SCADA system security architecture, Factor Neural Network Based SCADA Security and Defense Model [3] (FSDM) is constructed, as shown in Fig. 1.

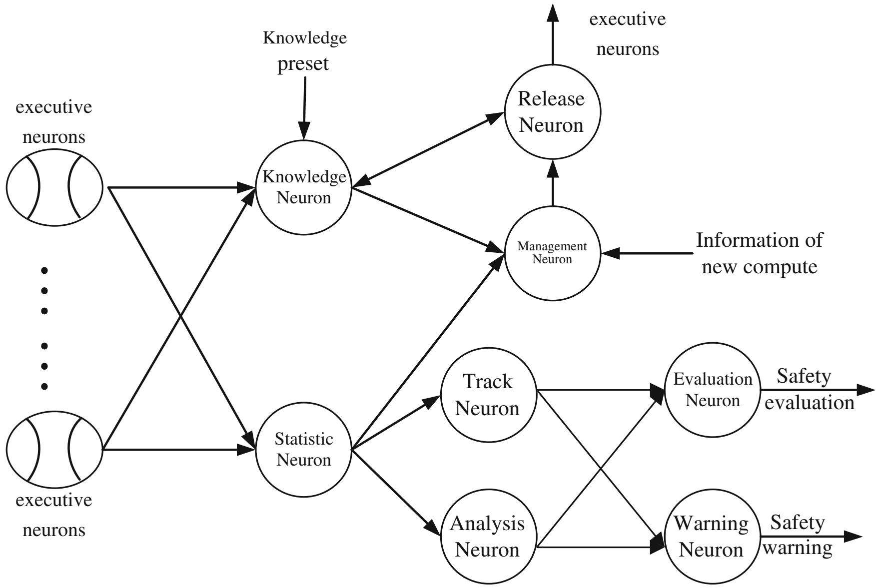


Fig. 1 Factor neural network based SCADA security and defense model

3 The Formal Method of Petri Nets

During The formal description for the SCADA model, macroscopically, establish Petri nets between neurons for the factor neural network of FSDM and describe the internal relationship; microscopically, describe the internal reasoning process of neurons by Fuzzy Petri Nets (FPN) theory.

3.1 The Basic Concept of Petri Nets

The Petri nets is a graphical mathematical modeling tool, it uses Connection of places, transitions and arc to express the system structure and function. Places express condition, resources and information which can be statically expressed in system. Transitions express the changes in system.

Definition 1 The Petri nets is a triple:

$$N = (P, T, F)$$

Among them, N represents places set; T represents transitions set and does not intersect with P; F represents relationship flow, $F \subseteq (P \times T) \cup (T \times P)$, namely arc set.

Definition 2 Fuzzy Petri Nets (FPN) is a sextuple:

$$FPN = (P, T, I, O, f, S_0)$$

Among them, (1) P = {p_1, p_2, …, p_n} is places nodes set, each element represents a fuzzy proposition;

(2) T = {t_1, t_2, …, t_m} is transitions nodes set and represents rule implementation conditions;
(3) I:P × T → [0, 1], represents fuzzy relation on P × T and weight coefficient of connection from places nodes to transitions nodes, and satisfy $0 \le I(p_i, t_j) \le 1$, i = 1, 2, … n, j = 1, 2, … m;
(4) O:T × P → [0, 1], represents fuzzy relation on T × P and confidence of connection from transitions nodes to places nodes, and satisfy $0 \le O(t_j, p_i) \le 1$, i = 1, 2, … n, j = 1, 2, … m;
(5) f is a real number of a value in [0, 1] defined in transitions nodes set T and represents trigger threshold of transitions nodes;
(6) S_0 is real number function of a value in [0, 1] on places nodes set P and represents flag state of the beginning when places nodes reasoning, namely initial confidence of proposition.

3.2 *Internal Reasoning Description of Neurons of FSDM*

3.2.1 Input Weight Coefficient of Fuzzy Petri Nets

Using the API HOOK technology to monitor malware API function calls, maliciousness of malicious program behavior factors is divided into four levels: low, medium, high and extremely high according to frequency of API function call when program running.

Maliciousness membership degree of malicious behavior factors can be determined by the membership function, the size of the membership degree can reflect essential characteristics of things correctly and intuitively. Different malicious behavior frequency follow certain distribution, the membership function for fuzzy set of malicious behavior factors can be constructed by normal fuzzy distribution [4], combined with data analysis for honeypot system, malicious behavior factor fuzzy set can be gained by fuzzy processing and quantifying maliciousness levels to X = {1, 2, 3, 4} (Corresponding to the four level of maliciousness: high, medium and low, extremely high).

As shown in Table 1.

According to membership table of Maliciousness level, malicious weight coefficient of malicious behavior factors can be calculated by weighting average decision method.

Table 1 Maliciousness membership of malicious behavior factors

Behavior factor	Membership for four level			
	1	2	3	4
TerminateProcess	0.1	0.8	0.1	0
SetWindowsHookEx	0.1	0.4	1.0	0.4
CreateService	0.2	0.5	0.2	0.1
StartService	0.1	0.8	0.1	0
DeleteService	0	0.2	0.6	0.2
ControlService	0.1	0.8	0.1	0
RegCreateKey	0.2	0.7	0.1	0
RegQueryValueEx	0.3	0.3	0.3	0.1
RegSetValueEx	0.1	0.3	0.4	0.2
SetFileTime	0.2	0.4	0.3	0.1
SetFileAttribute	0.3	0.3	0.4	0
CreateFile	0	0	0.9	0.1
FindNextFile	0.1	0.7	0.2	0

3.2.2 Formal Reasoning of Fuzzy Petri Nets

Definition 3 The matrix in the fuzzy Petri nets model can be defined as follows [5, 6]:

(1) $U = \{\mu_{ij}\}$ is the input matrix, $\mu_{ij} \in [0,1]$ is the weights from the places nodes P_i to the transitions $T_{j.}$. Among them, if the places nodes P_i is input nodes of transitions T_j, then μ_{ij} is equal to the weights of the directed arc $I(p_i, t_j)$ from P_i to T_j; Otherwise, then $\mu_{ij} = 0$, and $i = 1, 2, 3, \ldots, n$; $j = 1, 2, 3, \ldots, m$.

(2) $V = \{v_{ij}\}$ is the output matrix, $v_{ij} \in [0,1]$, it expresses that the confidence of the output directed arc from transitions T_j to the places nodes P_i. Among them, if the places nodes P_i is the output of transitions T_j, then v_{ij} is equal to the result confidence from transitions T_j to places nodes P_i; Otherwise, then $v_{ij} = 0$, and $i = 1, 2, 3, \ldots, n$; $j = 1, 2, 3, \ldots, m$。

(3) $S = [s_1, s_2, \ldots, s_n]^T$ is the state vector of the places nodes P, and represents the confidence of the corresponding propositions of places nodes, the initial confidence algorithm is gained by fuzzy processing of the factor object type.

(4) $F = [f_1, f_2, \ldots f_m]^T$ is the set of the transitions nodes threshold, f_j is corresponding triggering threshold of transitions T_j, and $f_j \in [0, 1]$, $j = 1, 2, 3, \ldots, m$.

To accomplish the definition of the above matrixes, two relational operators are also needed to define as following.

Definition 4 Set M、N、X、Y、Z are $n \times m$ matrix, then Comparative operator $\frown$: $X = M \frown N \Longleftrightarrow x_{ij} = 1, m_{ij} \geq n_{ij}; x_{ij} = 0, m_{ij} < n_{ij}$.

Multiplication operator $\oplus$:$Y = M \oplus N \Longleftrightarrow y_{ij} = m_{ij} \times n_{ij}$.

Max operator $\vee$: $Z = M \vee N \Longleftrightarrow z_{ij} = \max(m_{ij}, n_{ij})$ and: $i = 1, 2, 3, \ldots, n$; $j = 1, 2, 3, \ldots, m$.

There are N propositions and M reasoning rules in malware reasoning process of SCADA system host, which shows as N places nodes P and M transitions T in the FPN model, whose input matrix $U_{n\times m}$, output matrix $V_{n\times m}$, propositional confidence S_{1xn}, and the transitions threshold $F_{1\times m}$ can be obtained respectively according to the above definitions.

The reasoning algorithm can be described formally by matrix operation, which is divided into the following several steps:

(1) Calculate equivalent confidence Q when multi places nodes input

$$Q = U^T \cdot S_0$$

$Q = [q_1, q_2, \ldots, q_m]^T$ represents an input vector of weight coefficient is 1 which is transformed equivalently from initial confidence and Input weight coefficient of multi input places nodes of the same transitions.

(2) The comparison R between Equivalent input confidence and transitions threshold

$$R = Q \frown F$$

R represents whether the corresponding equivalent input can trigger the transitions threshold, $R = [r_1, r_2, \ldots, r_m]^T$, When the confidence q_j is greater than transitions threshold f_j, $r_j = 1$; Otherwise, $r_j = 0$, $j = 1, 2, 3, \ldots, m_o$

(3) Eliminate the input when equivalent confidence is less than transitions threshold

$$C = Q \oplus R$$

$C = [c_1, c_2, \ldots, c_m]^T$, Due to the elements in the R are only 1 and 0, so only the value which can trigger transition exist in the c, otherwise 0, the value indicates confidence of equivalent input places which can trigger transitions.

(4) The confidence of conclusion places outputted by the first reasoning in FPN.

$$S^1 = V \cdot C$$

S^1 is n-dimensional column vector and represents the confidence of the places after the first-round reasoning. The conclusion confidence is 0 when conclusion proposition places can not be gained by reasoning directly.

(5) Calculate the current confidence of all places nodes.

$$S_1 = S_0 \vee S^1$$

Set S^k the conclusion after the K-round reasoning, confidence of all the places nodes after the K-round iterative computation is $S_k = S_{k-1} \vee S^k$.

(6) After the first K-round reasoning, the confidence meets $S_k = S_{k-1}$, namely, after iterative computation, when the confidence of every places nodes does

not change, the reasoning process is completed. The output of the conclusion is the places of the highest confidence.

All the above derivation can be summarized to a formula, as follows:

$$S_k = S_{k-1} \vee (V \cdot C) = S_{k-1} \vee (V \cdot (Q \oplus R)) = S_{k-1} \vee (V \cdot (Q \oplus (Q \frown F)))$$
$$= S_{k-1} \vee V \cdot ((UT \cdot S_{k-1}) \oplus ((UT \cdot S_{k-1}) \frown F))$$

The conclusion confidence can be gained after iterative calculation when introducing initial confidence S_0.

Malicious program processing flow is shown in Fig. 2.

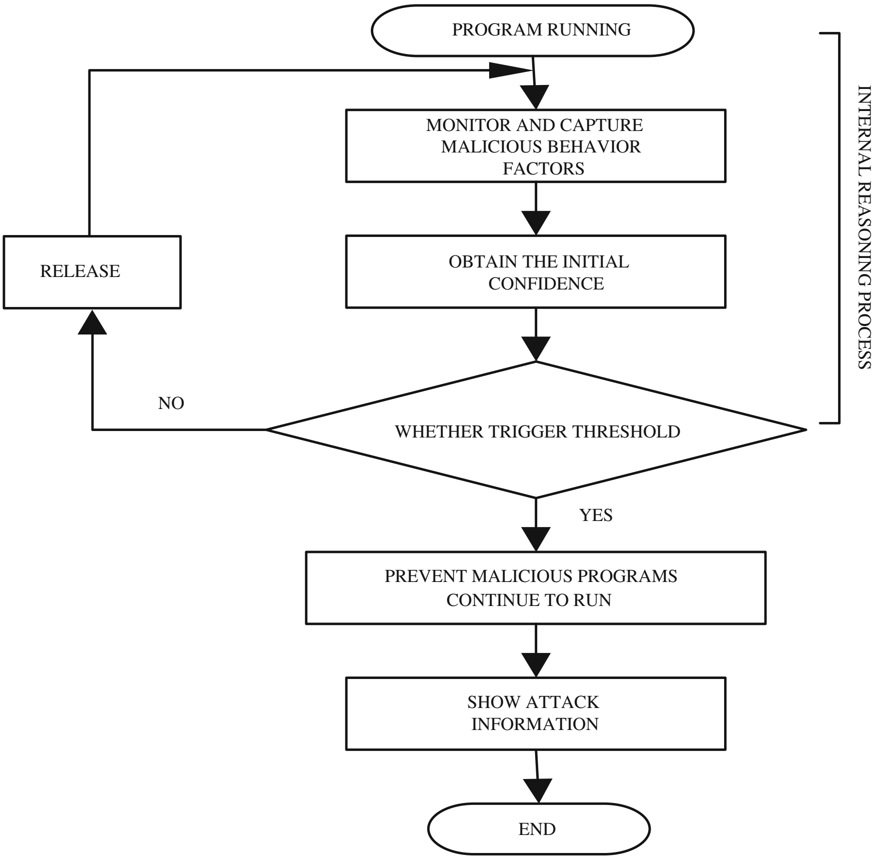


Fig. 2 Malicious program processing flow by executive neurons

3.2.3 Petri Nets Model of FSDM

SCADA security and defense factor neural network is a discrete event input-output system, the use of Petri nets discrete event system modeling, performance of the system needs to be mapped to the properties of Petri nets, Petri nets through performance analysis to achieve the actual system performance analysis The purpose of the establishment of Petri nets model FSDM structure, as shown in Fig. 3.

The places and transition of Petri nets expressed as shown in Table 2.

The SCADA security defense system running process:

Monitoring the behavior of the current running program by API monitoring technology, when found suspicious behavior, calculate equivalent confidence of a series behaviors of the program through the reasoning algorithm above. When the equivalent confidence is less than the threshold, the program should be continued to be released and monitored, and the malicious program wouldn't be identified through behavioral factors until the equivalent confidence is more than the threshold. After the malicious program is stopped, the current knowledge of related malicious behavior factors would be stored into the knowledge base, and after the corresponding rules of current behavior's factor knowledge is obtained, the rules would be managed and checked for updates, When the malicious program is blocked, the malicious behavior factors are being counted at the same time, and after obtaining the path information of malicious code, the attack source of the malicious program and related attack information would be tracking, then evaluating the safety of system after detection of malicious programs. At the same time, after the malicious behaviors' distribution and API function callings' distribution being counted, the malicious behavior factors would be analyzed, and it would give a warning prompt if there is a special malicious behavior.

Fig. 3 Petri nets model SCADA security defense system

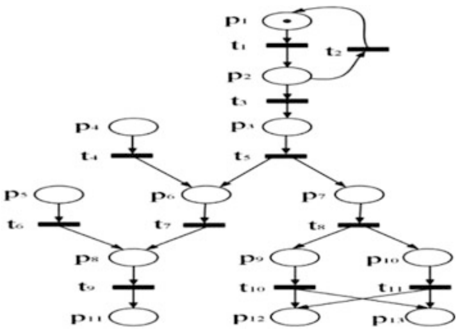


Table 2 Meaning of places and transitions

Places	Meaning	Transitions	Meaning
p_1	Program running	t_1	Monitor program behavior
p_2	Suspicious	t_2	Equivalent confidence < threshold
p_3	The status of receiving knowledge	t_3	Equivalent confidence > threshold
p_4	Identification of malicious programs	t_4	External joining knowledge
p_5	Computer configuration information	t_5	Stop malicious programs running
p_6	Knowledge base of malicious behavior factor	t_6	Computer for new join system
p_7	Statistics of malicious behavior factor	t_7	Obtain knowledge reasoning rules
p_8	Management state	t_8	Obtain distribution of malicious program behavior
p_9	Tracing the source of malicious programs	t_9	Other host information in the system is not updated
p_{10}	Analyze the current malicious program behavior factor	t_{10}	Uncovered malicious program source
p_{11}	The publication of the rule of reasoning and management	t_{11}	Existence of special malicious behavior
p_{12}	System security assessment		
p_{13}	Safe state warning alert		

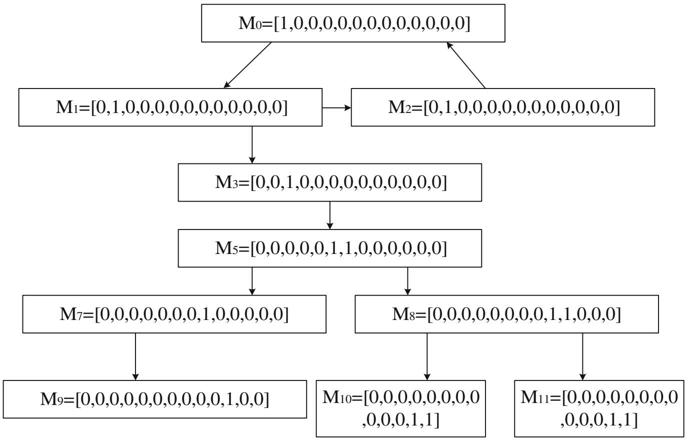


Fig. 4 Overlay tree of workflow

According to the running processes of the system, the workflow of Petri nets model in FSDM could be analyzed:

(1) There is a Token in the places p_1 in the initial state, and the corresponding transitions t_1 is active at this time. Only after the Token in p_1 is passed, then the state of the corresponding places p_1 and transitions t_1 would be changed. The state of token is $M = [m(p_1), m(p_2), \ldots, m(p_{13})]$, which means the Token state of all places in one state. The Token state in the initial state is $M_0 = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$.
(2) After passing the transition t_1, the state of places p_2 will change to 1, and the state of places will also change. At this point, the places in p_2 will selective trigger the transition t_2 and t_3, and only one of them two can be triggered. When the t_3 is triggered, the Token in p_2 will be transferred to place p_6 and p_7, the Token state in corresponding places at this point is $M_3 = [0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0]$. According to the analysis above, the overlay tree of workflow can be obtained, as shown in Fig. 4.

4 Conclusion

Petri nets is from the perspective of process, for the description and analysis of complex systems and design a model of effective tools, suitable for describing concurrent systems model, a formal steps and mathematical graph theory support theoretical rigor and technical characteristics for easy programming, based on FNN SCADA security defense system development laid the theory and technology.

Acknowledgment Thanks to the support by National Natural Science Foundation of China (No. 61175122).

References

1. Li, K.S., Ning, S.P.: Malicious script program and HOOK API based registry monitoring technology. Comput. Appl. **29**(12), 3197–3200 (2009)
2. Jia, X.Z., Guo, X.T., Zhang, H.S.: A virus detection method based on fuzzy pattern matching of program behavior. J. Qiingdao Univ. (NATURAL SCIENCE EDITION) **20**(4), 69–72 (2007)
3. Cao, X.D., Zhang, W.W., Hu, Q.C.: Research on SCADA security defense model based on FNN. Micro Comput. Appl. **34**(1), 10–12, 16 (2015)
4. Chen, L.: A Number of Key Technologies for Malicious Code Detection. Yangzhou University, Yangzhou (2012)
5. Xue, J.Y., Jia, L.X., Ru, F.: The fuzzy Petri net based formal inference algorithm and its application. J. Xi'an Jiao Tong Univ. **37**(12), 1263–1266 (2003)
6. Fan, X.M., Wu, R.H., Wu, J.: Forward inference algorithm for weighted fuzzy Petri nets. J. Dali Univ. **6**(8), 68–72, 81 (2007)

An Optimized Algorithm for Text Clustering Based on F-Space

Yu-Bin Zhong, Zhao-jun Li, Mao-han Zhao and Wei-quan Yao

Abstract In this paper, firstly, we express the text filter as a pretreatment of sequences of syllabic words by using the theory of factor space (F-space) to determine the basic factor set and then to do factor decomposition. Secondly, we structure text's factor matching vectors and their similarity degree, and then to achieve a deeper characterization of concepts by factors' similarity degree between texts, so as to realize the mining of text clustering. Finally, we randomly selected 90 articles on sogou.com as the experimental object, and verify the proposed algorithm by using plane division method and hierarchical agglomerative clustering these two algorithms. The results show that: (1) the clustering accuracies reaches 91 % and 94 % respectively; (2) the classification result obtained from proposed algorithm in this paper has little difference from the results of manual annotation method; (3) the effect based on hierarchical agglomerative clustering has better performance when compared with the plane division method. This paper provides a new feasible method for text mining fields.

Keywords F-space · Text clustering · Factor matching vector · Factors' similarity degree

1 Introduction

According to the Internet Development Statistics Report of China published by CINIC (China Internet Network Information Center) in January 2007, 70.2 % of the network information is presented in the form of plain text. For such semi-structured or unstructured data, finding approach to deriving the specific content of the

Y.-B. Zhong (✉) · Z. Li · M. Zhao
School of Mathematics and Information Sciences, Guangzhou University, Guangzhou 510006, China
e-mail: zhong_yb@163.com

W. Yao
Department of Basic Course, Guangzhou City Polytechnic, Guangzhou, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_44

information and knowledge from the text has become a problem. In recent years, text mining, data mining has become an important area of research [1]. Text mining enables us to compare documents, evaluate the importance of documents, calculate relevance between two documents, and identify writing patterns. However, since each text in the text set is composed by a large number of characteristics, such high dimension of text may significantly affect design and implementation of text mining system. Therefore processing text set with high-dimensional feature is obviously more sophisticated than traditional text mining system, the number of characteristics of natural language text and possible features Record number of combinations of values than we have found in relational databases or level database, the number of a lot more, even for the most modest text collections, the features collections used to represent the Chinese phrase level may still be extraordinarily complex, that the use of optimization feature class, characteristics and application of the concept of class-related fields reach the magnitude of thousands [1]. However, Wang Pei Zhuang proposed by the factor space [2, 3] erected a generalized coordinate system used to describe the actual object below the face of the text features of the “curse of dimensionality”, this paper presents a viable, high-performance method.

2 Preliminary

2.1 Factor

Factor as a meta word of factor space theory, can be portrayed from the following three aspects [4]:

(1) Attribution
When people get a good harvest, they want to consider the causes of the harvest, as abundant rainfall, so he attributes the rain as the main reason for the harvest. Factors state or its characteristics are different, the rainfall is a factor, 50, 100 ml and so are its characteristics. Factor is the public prompted various states and features associated therewith, the state is special tips about a factor, and the characteristic is about a factor of roughly prompt. Attribution has two meanings: one is the result of looking for reasons, this time caused by a factor understood as the result of something; the other is characterized by a state, this time factor will be used as a kind of status symbol or a set of features.

(2) Rationality
The formation process of concept is to find the difference between different things. For instance, the difference between “male” and “female” exists by comparing its meaning, while these two words have common factors. They are both used to represent the kind of sex. Factors can be understood as a way of understanding and modeling the real world, which can describe the

characteristic of the objective substance from different angles. The analysis process of text is the process of finding factors.

(3) Description
Every single object of the real world can be viewed as a set of descriptive factors. A person can be determined by his age, sex, height, weight and other performance. This intersection means you can establish a generalized coordinate frame; things are described as generalized coordinate system such a point. Establish key to this generalized coordinate system is to grasp, like age, sex and some other name, they are factors. Therefore, factor is generalized Dimension Name coordinate system.

2.2 State Space

If Factors u is associated with factor f, it means there is a corresponding state $f(u)$ compose a left pair with factor f, referred to as $(U, V]$, and if U and V consist of a number of factors and a set of objects, for any $u \in U$, all of the factors relevant to u are in V, Given a left pair $(U, V]$, you can define a relationship R between U and V: $R(u,f) = 1$, if and only if u is relevant to f.

We define $D(f)$, $V(u)$ as $\{u \in U | R(u,f) = 1\}$, $\{f \in V | R(u,f) = 1\}$. The factors $f \in V$ can be regarded as a mapping, when acting on a certain object $u \in U$ it will return a certain state $f(u)$:

$$f : D(f) \rightarrow X(f), u \mapsto f(u)$$

We define $X(f)$ mentioned above as $\{f(u) | u \in U\}$, and we call f the state space, any element in $X(f)$ is called a state of f. From the angle of state space, factors can be roughly divided into four types:

(1) Variable Factors: factors such as time, length, mass and other factors, the state space are generally one-dimensional or multi-dimensional Euclidean space or a space of its continuous or discrete subset.
(2) Symbolic Factors: such as occupation of this factor, the state space is composed of workers, farmers, teachers, lawyers and other nouns. They are symbols, representing the various things in the real world.
(3) Switching Factors: such as gender, the male and female composition of the two states, this factor can also be expressed as “M”, the state space by the Yes and No two states components.
(4) Degree Factors: such as the feasibility of such a class factor, creativity, satisfaction, there is no ready means of measurement, but there is a certain degree at all, its state space typically range [0, 1].

2.3 Relationships and Operation of Factors

(1) Zero factors. We introduce a special symbol θ, representing the empty state, In any state x, regardless of its composition and the composition of the set or sequence even do not work, that is $\{x,\theta\}=\{x\}$, $(x,\theta)=x=(\theta,x)$ we call symbol 0 zero factor, if $X(0)=\{\theta\}$, that is, zero factor in only one state, and the state or an empty state.

(2) Equal factors. We call factors f and factors g equal, if the mapping between factors f and factors g equals, that is $D(f)=D(g), X(f)=X(g)$, and for any $u\in D(f), f(u)=g(u)$.

(3) Sub-factors. In some cases, once factor A is confirmed, factor B state will be determined by factor A. If factor f represents the coordinate of the point on the plain. Factor g represents point of the horizontal factor, the state (position points (x, y)) of f will determine the g 's state (location of the point x). In this relationship, the state space B may represent a subspace A of the state space. Factors g is called the factor f's true subfactor, denoted by $f > g$, if there is a collection Y has $Y\neq\varnothing$ and $Y\neq\{\theta\}$, so $X(f)=X(g)\times Y$, then we called g the sub-factors of f, referred to as $f\geq g$, if $f>g$ or $f=g$.

(4) The combined factors. We call factors h the combined factors of factors f and factors g, denoted by $h=f\wedge g$. If h is the largest common sub-factors, that is, and for any one factor e, there must be $f\geq e$, $g\geq e\Rightarrow h\geq e$. Factors g is called the combined factors of factor family $\{f\}_{(t\in T)}$, recorded as $g=\bigwedge_{t\in T} f_t$. If g is the greatest common sub-factors of $\{f\}_{(t\in T)}$, that is $\forall t\in T, f_t\geq g$, and for any factors h that must be $\forall t\in T, f_t\geq h\Rightarrow g\geq h$.

(5) Factors disjunction. Factors h is called the disjunction factor of factors f and factors g denoted by $h=f\vee g$. If $h=f\vee g$ is a sub-factors of f, g, and it is the smallest of such factors, that is $h\geq f$, $h\geq g$, and for any factors e that must be $e\geq f, e\geq g\Rightarrow e\geq h$. We call Factors g as disjunction factor of the factor family $\{f\}_{(t\in T)}$, referred to as, $g=\ \vee_{t\in T} f_t$ if that sub-factors, and is the smallest of such factors, that is $\forall t\in T, g_t\geq f$, and for any factors h that must be $\forall t\in T, h\geq f_t\Rightarrow h\geq g$.

(6) Independent factors. We call factor family $\{f_t\}_{(t\in T)}$ is pairwise independent, if the condition $\forall s,t\in T, f_s\wedge f_t=0$ is satisfied.

(7) Difference: Define factors h as the difference between f and g, referred to as $h=f-g$, if $(f\wedge g)\vee h=f, h\wedge g=0$ is satisfied.

(8) The remainder of factor: In a problem, we often consider only one class of factors F relevant to the issue. We Call 1 the whole factor of factor F. If $1\in F$ and $\forall f\in F, 1\geq f$ to any one factor f, denoted $1-f$ by f^c. If $f^c\in F$, then called f^c the remainder of factor 1.

(9) Atomic factor. Factor f is called atomic factor, if in addition to the zero factor, there is no real sub-factors for the above-mentioned factors for a class, the all atomic factor referred to as π. F denotes the atomic factor family.

3 Text Clustering Algorithm Based on Factors Space

Text mining algorithms text pre-processing operations will transform non-standard text and implicit structure into defined structure. The aim of the process is to identify a subset of simple textual features, which will be used to indicate the specific text. Word, phrase, short sentence and concepts are the four most commonly used text feature. Although conceptual stage method depends on human interaction, the representation of concept performs much better than other method when dealing with synonyms, polysemy. By using the body and domain knowledge repository provided representation based on the concept can be further used to support highly complex conceptual level, and to provide the best representation.

Therefore by decomposing the concept based on the factor space theory we can attain a deeper characterization of the concept, which is a fundamental step of text clustering.

Text-based spatial clustering algorithm factors include the following stages, which can be briefly described as follows.

3.1 Text Preprocessing

Text preprocessing stage includes Chinese word segmentation, old-fashion words filtered.

Chinese word segmentation is cutting some Chinese character sequence into word sequence of process Chinese word with some segmentation method or algorithm. Segmentation algorithms can be broadly divided into three categories: dictionary segmentation, statistical segmentation and rule segmentation [5]. By comparing the three algorithms, we can easily find that dictionary segmentation do not need the rule base as support; furthermore, dictionary word is the most mature technology; moreover, dictionary segmentation algorithm has an advantage over others from the angle of algorithm complexity and speed, but it's performance drops conspicuously when dealing with ambiguity recognition and new word recognition. Because the experimental material herein Sogou corpus of news, and news text are generally more standardized, and the identification of new words from the dictionary make up the choice, so we chose the dictionary word segmentation.

After completing the segmentation of the text, we obtained words and punctuation character sequence. In this sequence of characters, not only the existence of various parts of speech of words, and various punctuation marks. But for some vocabulary and punctuation, and their existence are not only the meaning of the sentence did not have much impact, but will greatly increase the computing capacity of the system. Therefore, the results need to be filtered word processing, including removal of function words such as “是”, “的” and “和”, and the removal of the short phases with low semantic contribution, such as “某种”, “各种”, “能够”, “相当于, “还有一点”, “称之为”and so on.

3.2 Determination of the Basic Factors Set

Determination process of the basic factors includes factors extraction, factors factorization and the establishment of basic factors set.

After the text preprocessing, we attain a series of word sequence. Then we need to further extract factors on the basis of these terms of a sequence. Firstly, we should extract high-frequency words, word sequences for preliminary screening in high-frequency word further according to artificial understanding, combined with domain knowledge, to establish a set of initial factors such as extraction and education-related factors texts extracted human factors, its state space for students, teachers, and candidates.

For different topics of text, the state space is different. As a group talk extracting text relating to these two types of education or sports, are able to extract factors personnel, but its value of state space for students, teachers, candidates, players, coaches, etc., of which the first three states are talking about education, and then the two states are talking about sports, even if the factor of the same name, its state space are not the same, so we need to decompose the human factors into education—human factors and Sports—the human factor two factors.

After the above two steps, we obtain the initial factor set in the initial factor concentrate, the ability to characterize the entire text of each of these factors on the set is different for a factor, if multiple text are in the dimension of the value of this factor 1 indicates that the capacity of the factors characterizing the entire text of the collection is relatively strong, if only a few values of text projected on the dimension of this factor is one of the factors that indicate the ability to characterize the text set is relatively weak. In order to simplify the complexity of the subsequent calculations, the removal of the initial factors focus on the overall text characterizing weak factors, such factors resulting set is called set of fundamental factors calculated similarity factor in the subsequent text are set to perform the basic factors.

3.3 The Calculation of Text Factor Similarity

The structure of the vector is an essential element set of text. So we can turn the problem of evaluating similarity between the unstructured text into evaluating the similarity between the structure of the vector.

According to different values of the factors match vector circumstances, can be divided into the following four cases were assigned factors match vector on one factor dimension d:

(1) If two text values in the dimension d are nonzero and equal, then factor in the match vector values in dimension d is 1;
(2) If two text values are zero in on dimension d, the value of factor in the match vector in the dimension d is β;

(3) If the two text values in the dimension d are not equal, but the intersection of the two values is not an empty set, then the value of factor in the match vector in the dimension d is α;
(4) If the two text values in the dimension d are not equal, but the intersection of the two values is the empty set, the value of factor in the match vector in the dimension d is 0.

The text between the degrees of similarity between the factors text mainly takes nonzero vector in dimension matching factor. Factor matching vector value of dimension 1, the value of b value of a dimension and a dimension Influencing factors are not the same as the similarity of common values are calculated at the time, should be given different weights of its common greatest impact on the value of the dimension is 1, the maximum weight should be assigned; on common minimal impact is the value of dimension β, its weight assignment should be minimal; the impact value of dimension α of the factors in the similarity between the first two, but also for its weight assigned value is set between the first two dimension 1 half of the total weight of the dimensions, ranging from the right to re-dimension β is 1, and the value of a weight is dimension α of the total value of between half and one. Thus, the similarity factor equals to match vector elements take non-zero values in all the dimensions and weights of the accumulation to obtain the following formula [5]:

$$x(T_1, T_2) = c_1 * \frac{D}{2} + c_\beta + c_\alpha * \lambda \tag{1}$$

$x(T_1, T_2)$ represents the similarity between the factors text T_1, T_2; c_1 represents factor matching vector value of dimension number 1, c_β indicates factor matching vector value β is the number of dimensions, c_α shows the value of the factors match vector dimension α the number of factors that match vector. D indicates the total number of dimensions, that value λ is the weight corresponding to the weight factor, its value between 1 and $D/2$. At the time of the similarity factor calculation, using the calculated values obtained by the method described above is greater than 1 values for subsequent conversion to facilitate the processing of the distance, the similarity factor should be between 0 and 1 values. Therefore, the use of normalization formula factors similarity is normalized to ensure that its value at 0–1 between normalized formula is as follows:

$$c(T_1, T_2) = \frac{2}{1 + k^{-x(T_1, T_2)}} - 1 \tag{2}$$

$c(T_1, T_2)$ represents factor normalized similarity; $x(T_1, T_2)$ is the results of formula (1). Formula (2) is the normalized formula, it can normalize values in the range of [0, 1] into values between [0, 1], which is set to a number greater than 1 final, 1 minus the similarity factor normalized to obtain factor between distance formula text as follows:

$$d(T_1, T_2) = 1 - c(T_1, T_2) \tag{3}$$

wherein, $d(T_1, T_2)$ represents the distance between the text element; $c(T_1, T_2)$ is the calculation results obtained by the formula (2).

After deduction available, some properties of factor similarity and distance factors:

Property 1 *For any text T, when $T_1 = T$ the similarity factor $x(T, T_1)$ reach the maximum value.*

Proof According to the formula for calculating the similarity factor formula (1), for different values of text T_1 on a dimension, it certainly has the formula (1) is an c_1, c_α, c_β different values, which have $x(T, T_1) \leq x(T, T)$.

When $x(T_1, T_2)$ get maximum value $(d(T_1, T_2)$ get minimum value), according to Property 1, when $T_1 = T_2$ factor when taking the minimum distance $d(T_1, T_2)$, so may wish to provide as $T_1 = T_2$, apparently arbitrary text $d(T_1, T_2) = 0$, then have the following..:

Property 1′ *For any text T_1, T_2, are; there were $d(T_1, T_2) \geq 0$. When $T_1 = T_2$, we have $d(T_1, T_2) = 0$.*

Property 2 *For any text T_1, T_2, we have $d(T_1, T_2) = d(T_2, T_1)$.*

Proof The factor of the similarity calculation Chinese position is symmetrical, so we have $x(T_1, T_2) = x(T_2, T_1)$, thereby normalizing factor similarity $c(T_1, T_2) = c(T_2, T_1)$, so $d(T_1, T_2) = 1 - c(T_1, T_2) = 1 - c(T_2, T_1) = d(T_2, T_1)$.

Property 3 *For any text T_1, T_2, T_3, there is $d(T_1, T_2) \leq d(T_1, T_3) + d(T_3, T_2)$*

Proof To prove $d(T_1, T_2) \leq d(T_1, T_3) + d(T_3, T_2)$, we need to prove $1 + c(T_1, T_2) \geq c(T_1, T_3) + c(T_3, T_2)$, according to Eq. (2), expand and simplify this formula, then we can get

$$2k^{x(T_1, T_2)} + 1 + k^{x(T_3, T_2) + x(T_1, T_2)} + k^{x(T_1, T_3) + x(T_1, T_2)} \geq k^{x(T_1, T_3) + x(T_3, T_2)}$$

According to the basic inequality has

$$2k^{x(T_1, T_2)} + 1 + k^{x(T_3, T_2) + x(T_1, T_2)} + k^{x(T_1, T_3) + x(T_1, T_2)} \geq 5k^{\frac{4}{5}x(T_1, T_2) + \frac{1}{5}(x(T_1, T_3) + x(T_3, T_2))}$$

So just permit it as long as

$$x(T_1, T_2) - x(T_1, T_3) - x(T_3, T_2) \geq \frac{5}{4} \log_k \frac{1}{5}$$

Since $x(T_1, T_2)$ is bounded, it is possible to select the appropriate k to hold the above formula for any text T_1, T_2, T_3.

The above Properties of 1–3 indicate that if we add factors distance $d(T_1, T_2)$ between the texts, the text set meets the definition of distance:

(1) Symmetry: For any text T_1, T_2, there is $d(T_1, T_2) = d(T_2, T_1)$
(2) Triangle inequality: For any text T_1, T_2, T_3, there is $d(T_1, T_2) \leq d(T_1, T_3) + d(T_3, T_2)$.

This explains that the distance $d(T_1, T_2)$ is a suitable value for the measurement of the distance between the texts.

4 Experiment Verification

This article uses Sougou text classification corpus as experimental material. Sogou corpus from Sohu news website to save a lot of classified information news corpus and corresponding edited manually sorting and classification. Classification system includes dozens of classified nodes, the size of about ten pages million articles in document [6]. This paper select one of three categories: education, finance, sports. We randomly selected 30 articles from each class, a total of 90 articles and use of this algorithm for text clustering.

Text clustering algorithm based on the distance of the main plane division method and the level of cohesion and French methods, this paper uses two text clustering algorithm based on the distance of the experiment, the results shown in Fig. 1 two dendrogram from Fig. 1 can be found, 90 texts in both approaches can be broadly clustered into three broad categories. clustering results were examined every one of the text under two algorithms, the findings text clustering algorithm based on planar partition method for the 27 texts and two Financial Education gathered into a large class of text, 28 text and 3 educational texts gathered into a sports categories, 27 sports and three Financial text gathered into a category; and based on the results of coagulation process hierarchical clustering algorithm for text 30 text and 2 Financial Education text gathered into one category, 27 educational texts and two sports gathered into a large class of text, text with 28 sports an educational text gathered into one category thus learn the results clustering algorithms and manual classification results roughly the same.

Zhou [7] has proposed Class_F value on the basis of artificial label classes as a method of evaluation for text clustering. For each artificially defined topic P_j among the data, assuming that there is a group C_i in the hierarchy structure formed by the clustering algorithm. In order to find C_i, traverses all clustering in the results set $C = \{C_1, C_2, \ldots, C_m\}$, calculate its accuracy, denoted as *Precision*, recall ratio and F value, select the optimal index value and the corresponding cluster [8], and regard it as the best quality indicator to determine the quality of P_j.for any artificially defined themes P_j and clustering cluster C_i:

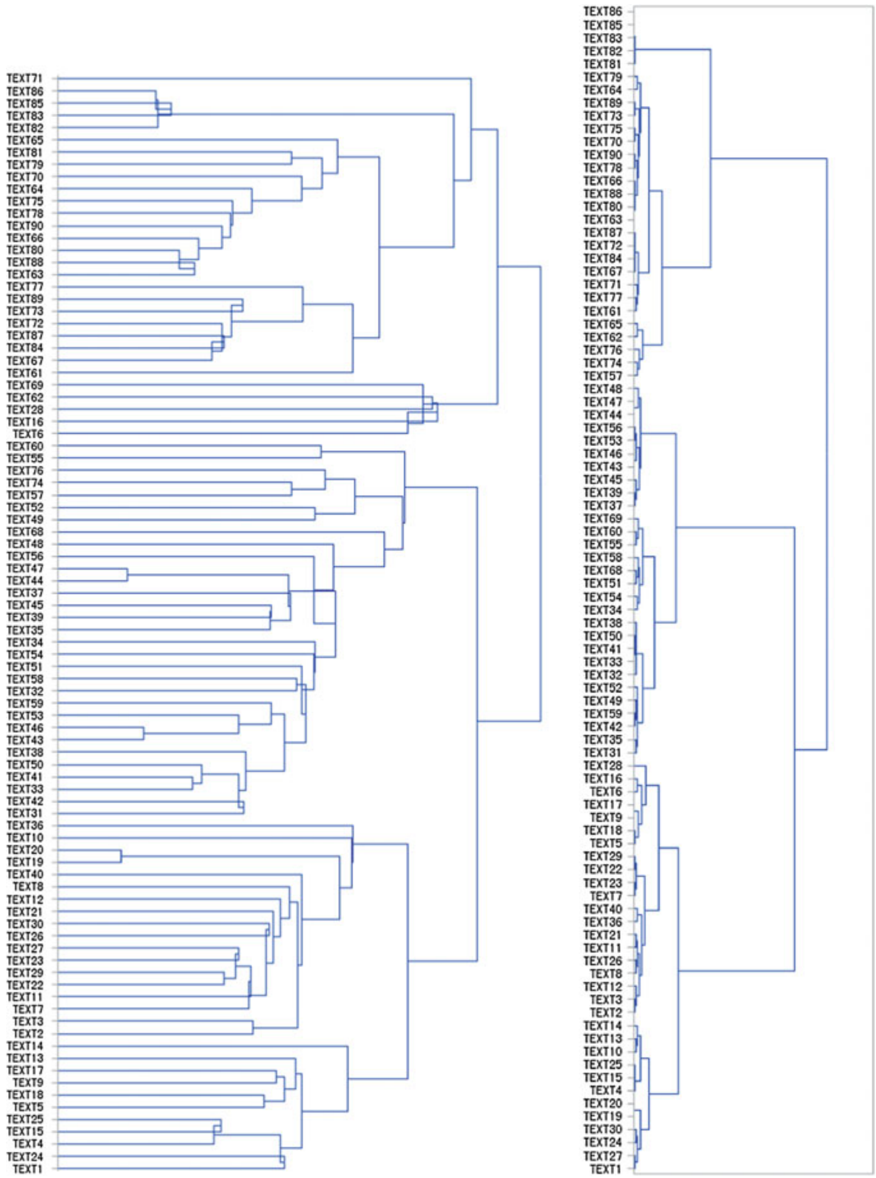


Fig. 1 Text clustering tree (*left* using planar partition method; *right* using hierarchical agglomeration method)

$$\text{Precision}(P_j, C_i) = \frac{|P_j \cap C_i|}{|C_i|} \tag{4}$$

$$\text{Recall}(P_j, C_i) = \frac{|P_j \cap C_i|}{|P_j|} \tag{5}$$

$$F(P_j, C_i) = \frac{2P(P_j, C_i)*R(P_j, C_i)}{P(P_j, C_i) + R(P_j, C_i)} \tag{6}$$

For each artificial theme P_j:

$$F(P_j) = \max_{i=1,2,\cdots,m} \{F(P_j, C_i)\} \tag{7}$$

The value of Class_F:

$$F = \frac{\sum_{j=1}^{s} |P_j|*F(P_j)}{\sum_{j=1}^{s} |P_j|} \tag{8}$$

With the above evaluation methods, we calculated the values of Class_F of two different text clustering algorithms.

Finally, by using text clustering algorithm based on the calculated plane division method [9], we calculated the value of Class_F is 0.911095675, when using algorithm based on hierarchical clustering, the result of Class_F is 0.94405, which indicates both two clustering algorithms can perform with high accuracy (more than 90 %), The result of calculation demonstrate the effectiveness of the algorithm. In addition, from Class_F values of the two algorithms, the effect of text-based coagulation process hierarchical clustering algorithm works better than division method based on plane [10, 11] (Table 1).

Table 1 F value of each class

Artificial label classes	Plane division algorithm	Hierarchical clustering algorithm
P_1	0.915254237	0.967741935
P_2	0.918032787	0.915254237
P_3	0.90	0.949152542

5 Conclusion

In this paper, we propose a feasible idea which applies the factor space theory to the field of text mining [12]. In subsequent work, we will continue improving the algorithm repeatedly in order to achieve better clustering and practical effect.

Acknowledgements This research was supported by Information and Computing Science Outstanding Talent Training project of Guangdong Province. (No.20153324)

References

1. Cheng, X.Y., Zhu, Q.: Text Mining Principle. Science Press, Beijing (2010)
2. Wang, P.Z., Sugeno, M.: The factors field and background structure for fuzzy subsets. Fuzzy Math. **2**(2), 45–54 (1982)
3. Zhuang, W.P., Li, H.X.: Mathematical Theory of Knowledge Representation. Tianjin Science and Technology Press, Tianjin (1994)
4. Li, H.X.: Mathematical framework factor space theory and knowledge representation—factor space frame axiomatic definition and description. Beijing Normal Univ. (Nat. Sci.) (1996)
5. Liu, Y.: Implicit policy kinship mining domain-oriented Harbin. Harbin Eng. Univ. (2013)
6. Sogou Sogou laboratory data download - text classification corpus. http://www.sogou.com/labs/dl/c.html
7. Zhou, Z.T.: Text Clustering Analysis Evaluation and Text Representation. Beijing, Chinese Academy of Sciences (2005)
8. Wang, P.Z.: A factor spaces approach to knowledge representation. Fuzzy Sets Syst. (1990)
9. Luo, C.Z., Yu, F.S.: Mathematical models and expert system development tool. Fuzzy Syst. Math. **6**(2), 20 (1992)
10. He, Q., Tong, Z.M.: A method of forming the concept of factors of space and fuzzy clustering. Syst. Eng. Theory Pract. (1999)
11. Liu, Y.M.: Feature extraction and classification factor space. Beijing Normal Univ. (Nat. Sci.) **36** (2), 172–177 (2000)
12. Wang, P.Z., Li, H.X.: Fuzzy System Theory and Fuzzy Computer. Science Press, Beijing (1996)

Part V
Information Granulation and Granular Computing

Parallel Quantum-Behaved Particle Swarm Optimization Algorithm with Neighborhood Search

Hai-xia Long, Man-zhi Li and Hai-yan Fu

Abstract In order to escape from premature convergence and improve the efficiency of the quantum-behaved particle swarm optimization (QPSO) algorithm, this paper propose a new algorithm PNSQPSO, which employing neighborhood search strategies (local search strategies and global search strategies) into QPSO to increase the diversity of population and parallel technique to shorten the running time of algorithm. A comprehensive experimental study is conducted on a set of benchmark functions, Comparison results show that PNSQPSO obtains a promising performance and less time cost on the majority of the test problems.

Keywords Quantum-behaved particle swarm optimization algorithm · Parallel · Local search strategies · Global search strategies · Premature convergence · Efficiency

1 Introduction

Particle swarm optimization (PSO) is a kind of stochastic optimization algorithms proposed by Kennedy and Eberhart [1] that can be easily implemented and is computationally inexpensive. The core of PSO is based on an analogy of the social behavior of flocks of birds when they search for food. PSO has been proved to be an efficient approach for many continuous global optimization problems. However, as demonstrated by Van den Bergh [2], PSO is not a global convergence guaranteed algorithm because the particle is restricted to a finite sampling space for each of the iterations. This restriction weakens the global search ability of the algorithm and may lead to premature convergence in many cases.

Recently, a new variant of PSO called quantum-behaved particle swarm optimization (QPSO) [3], which is inspired by quantum mechanics and particle swarm

H. Long · M. Li (✉) · H. Fu
School of Information Science Technology, Hainan Normal University,
Haikou 571158, China
e-mail: lmz20031979@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_45

optimization model. QPSO has only the position vector without velocity, so it is simpler than standard particle swarm optimization algorithm. Furthermore, several benchmark test functions show that QPSO performs better than standard particle swarm optimization algorithm. Although the QPSO algorithm is a promising algorithm for the optimization problems, like other evolutionary algorithm, QPSO also confronts the problem of premature convergence, and decrease the diversity in the latter period of the search. Therefore a lot of revised QPSO algorithms have been proposed since the QPSO had emerged. Differential mutation operation was adopted to enhance the global search ability in QPSO [4]. In Sun et al. [5], the mechanism of Gaussian distribution was proposed to make the swarm more efficient in global search. In Ref. [6], Niu combined QPSO with a selective probability operator to solve the economic dispatch (ED) problems with valve-point effects and multiple fuel options. Sun et al. [7] proposed a new scheme for clustering gene expression datasets based on a modified version of Quantum-behaved Particle Swarm Optimization (QPSO) algorithm, known as the Multi-Elitist QPSO (MEQPSO) model.

Like many other evolutionary algorithms, the main problems of QPSO face is how to enhance the search ability of particle, in order to prevent the particles fall into local optimal solution resulted in search ability drop. In this paper, QPSO with neighborhood search strategies including local search strategies and global search strategies) are introduced. This strategy is to prevent the diversity of particle swarm declining in the search of later stage. QPSO needs to solve another problem is to shorten the time cost. Along with the increase of size of optimization problem, if the algorithm has a good performance, then the population size must be increased, and then making the algorithm of large quantities of computation. PNSQPSO algorithm is proposed to solve the problem. In order to overcome this shortcoming, increasing the practicability of QPSO, combined with the development of computer technology, this paper introduce the parallel technology, so as to shorten the calculation time and improve efficiency.

2 QPSO Algorithm

In the original PSO with M individuals, each individual is treated as an infinitesimal particle in the D-dimensional space, with the position vector and velocity vector of particle i, $X_i(t) = (X_{i1}(t), X_{i2}(t), \ldots, X_{iD}(t))$ and $V_i(t) = (V_{i1}(t), V_{i2}(t), \ldots, V_{iD}(t))$. The particle moves according to the following equations:

$$V_{ij}(t+1) = V_{ij}(t) + c_1 \cdot r_1 \cdot (P_{ij}(t) - X_{ij}(t)) + c_2 \cdot r_2 \cdot (P_{gj}(t) - X_{ij}(t)) \quad (1)$$

$$X_{ij}(t+1) = X_{ij}(t) + V_{ij}(t+1) \quad (2)$$

for $i = 1, 2, \ldots M$; $j = 1, 2 \ldots, D$. The parameters c_1 and c_2 are called the acceleration coefficients. Vector $P_i = (P_{i1}, P_{i2}, \ldots, P_{iD})$ known as the personal best position, is the best previous position (the position giving the best fitness value so far) of particle i; vector $P_g = (P_{g1}, P_{g2}, \ldots, P_{gD})$ is the position of the best particle among all the particles and is known as the global best position. The parameters r_1 and r_2 are two random numbers distributed uniformly in (0, 1), that is $r_1, r_2 \sim U(0,1)$. Generally, the value of V_{ij} is restricted in the interval $[-V_{\max}, V_{\max}]$.

The PSO algorithm may be achieved if each particle converges to its local attractor $p_i = (p_{i1}, p_{i2}, \ldots p_{iD})$ with coordinates

$$p_{ij}(t) = (c_1 r_1 P_{ij}(t) + c_2 r_2 P_{gj}(t)) / (c_1 r_1 + c_2 r_2), \text{or}\, p_{ij}(t) = \varphi \cdot P_{ij}(t) + (1 - \varphi) \cdot P_{gj}(t) \tag{3}$$

where $\varphi = c_1 r_1 / (c_1 r_1 + c_2 r_2)$. It can be seen that the local attractor is a stochastic attractor of particle i that lies in a hyper-rectangle with P_i and P_g being two ends of its diagonal.

In Sun et al. [3] a global point called Mainstream Thought or Mean Best Position of the population is introduced into PSO. The mean best position, denoted as C, is defined as the mean of the personal best positions among all particles. That is

$$\begin{aligned} C(t) &= (C_1(t), C_2(t), \ldots, C_D(t)) \\ &= \left(\frac{1}{M} \sum_{i=1}^{M} P_{i1}(t), \quad \frac{1}{M} \sum_{i=1}^{M} P_{i2}(t), \ldots, \frac{1}{M} \sum_{i=1}^{M} P_{iD}(t) \right) \end{aligned}$$

where M is the population size and P_i is the personal best position of particle i. Then the value of L is evaluated by $L = 2\alpha \cdot \left| C_j(t) - X_{ij}(t) \right|$ and the position are updated by

$$X_{ij}(t+1) = p_{ij}(t) \pm \alpha \cdot \left| C_j(t) - X_{ij}(t) \right| \cdot \ln(1/u) \tag{5}$$

where parameter α is called Contraction-Expansion (CE) Coefficient, which can be tuned to control the convergence speed of the algorithms. In most cases, α decrease linearly can be controlled to α_0 to α_1 ($\alpha_0 < \alpha_1$) [3].

3 Parallel Quantum-Behaved Particle Swarm Optimization Algorithm with Neighborhood Search (PNSQPSO)

3.1 QPSO with Neighborhood Search (NSQPSO)

For each particle, its neighborhood may cover better solutions. To improve the ability of exploitation, a local neighborhood search strategy is proposed. During searching the neighborhood of a particle P_i, a trial particle LX_i is generated as follows [8]:

$$LX_i = r_1 \cdot X_i + r_2 \cdot pbest_i + r_3 \cdot (X_c - X_d) \quad (6)$$

where X_i is the position vector of the ith particle, $pbest_i$ is the previous best particle of P_i, X_c and X_d are the position vectors of two random particles in the k-neighborhood radius of $P_i, c, d \in [i-k, i+k] \wedge c \neq d \neq i, r_1, r_2$ and r_3 are three uniform random numbers within (0, 1), and $r_1 + r_2 + r_3 = 1$. The random numbers r_1, r_2 and r_3 are the same for all $j = 1, 2, \ldots, D$, and they are generated a new in each generation.

Besides the local neighborhood search, a global neighborhood search (GNS) strategy is proposed to enhance the ability of exploration. When search the neighborhood of a particle P_i, another trial particle $G_i = (GX_i, GV_i)G_i$ is generated as follows [8]:

$$GX_i = r_4 \cdot X_i + r_5 \cdot gbrst + r_6 \cdot (X_e - X_f) \quad (7)$$

where $gbest$ is the global best particle, X_e and X_f are the position vectors of two random particles chosen for the entire swarm, $e, f \in [1, N] \wedge e \neq f \neq i, r_4, r_5$, and r_6 are three uniform random numbers within (0,1), and $r_4 + r_5 + r_6 = 1$. The random numbers r_4, r_5 and r_6 are the same for all $j = 1, 2, \ldots, D$, and they are generated a new in each generation. The GNS strategy is helpful to solve multimodal problems. Particles are located at different regions. Therefore, if the current particle falls into local minima, particles in other regions may pull the trapped particle forward.

In every generation, if the particle's current position is better than its parent then replaced with current particle position; otherwise, we keep parent particle unchanged. After this operation, the local search strategy is conducted with p_{ns} probability. If the probability p_{ns} is satisfied, two particle LX_i and GX_i are generated. Then, the fittest particle among X_i, LX_i and GX_i is selected as the new P_i.

Procedure of NSQPSO:

Step 1 Initialize particles with random position; set the *pbest* position and *gbest* position of each particle;
Step 2 For n = 1 to $n_{\max}$ (maximum number of iterations), execute the following steps;
Step 3 Calculate the mean best position among the particles according to Eq. (4);
Step 4 For each particle, execute Step 5–7;
Step 5 Computer its objective function $f(X_{i,n})$. If $f(X_{i,n}) < f(P_{i,n})$ then $P_{i,n} = X_{i,n}$;
Step 6 Select the current *gbest* position;
Step 7 if *rand* (0, 1) $\leq p_{ns}$ then generate a trial particle LX_i and GX_i according to Eqs. (6) and (7).

Step 8 Calculate the objective function, $f(X_{i,n})f(LX_{i,n})$ and $f(GX_{i,n})$.
Step 9 Select the fittest one among X_i, LX_i and GX_i;
Step 10 Update *pbest* position and *gbest* position;
Step 11 For each dimension of each particle, get the stochastic position P by Eq. (3).
Step 12 Update each component of the current position by Eq. (5) and return to Step 2.

3.2 NSQPSO with Parallel Technique (PNSQPSO)

With the development of the microprocessor technology and network technology, reducing the computational cost becomes the possibility by the cluster, in order to make use of these technologies, the communication layer such as MPI and PVM has been used to optimize the design of parallel development, the QPSO algorithm is a new model of particle swarm optimization based on quantum theory derived, due to its inherent parallelism implied, so it can be easily parallelized, and there is no need to adjust many parameters, but also for most problems, parameter setting can complete the general optimization better. Introducing parallel mechanism of QPSO algorithm, which not only optimal QPSO algorithm on the value has been significantly improved, but also saves the running time of CPU, the node and the algorithm is not limited to the number of computer. PDCQPSO is combined high speed parallel with randomness of QPSO algorithm, greatly improving the speed and quality of QPSO algorithm.

In this paper, parallelization of PDCQPSO algorithm is used peer-to-peer mode, its core idea is: a group is divided into q sub group, q is the number of PPEs (Physical Processing Elements), PPEs regularly exchange of *gbest* and are synchronous communication, in order to all of the value of *gbest* broadcasting to all PPEs. Sub population storage *gbest* value obtained from other correspond all sub groups, and selected a *gbest* value to replace the current *gbest* from random, and in accordance with the (5) adjusts its position.

Period of PPEs communications is set to N and N is an exponential decline order. Its value is $\left|N_g/2\right|, \left|N_g/4\right|, \left|N_g/8\right|, \ldots$, and so on. Where N_g is a maximum number of iterations in the main loop. The principle is: at the start of the search, the diversity of the whole group is higher, local search is more important than global search, so PPEs should be in the sub groups independently working long time. When the search reaches the later stage, it is possible that every processor particle convergence to a different *gbest* value, then the global search is more important than local search. In the course of being optimized local *gbest*, this is necessary to continue to search in the domain of promising position, and avoid unnecessary work. The exchange the value of *gbest* is to increase the particle search diversity and obtain better global search.

The basic starting point is developing parallelization of population evolution from a global perspective. The population distribution in the local memory of each processor, then individual of each processor is independently optimized by QPSO, so there is linear speedup. The technique can effectively improve the computing speed of QPSO and maintain swarm local characteristics of the sub processor, so it can effectively avoid premature convergence phenomenon.

The t generation sub swarm of the P_i processor is denoted by $P_i(t)$. PDCQPSO algorithm running on the P_i processor is described using pseudo code in the following.

4 Test Environment and Test Function

In the simulation study, the PNSQPSO algorithm is implemented in a cluster environment, because MPI provides a collection of operation in group communications, here we use of communication patterns of peer to peer to realize PNSQPSO algorithm. Inter process call blocking for synchronization before collection calculation. Cluster environment configuration is as follows: a total of eight nodes, each node is a single CPU structure, use the CPU Inter Core 2.5 GHz, 4 GB RAM. Communication network is a LAN communication system of the 100 M and the star topology structure. Operating system is Windows XP. The communication environment is MPI. Version specific MPICH—2 is adopted. Using the VC bundled with MPI function interface is to be programming.

To test the performance of the PNSQPSO, five widely known benchmark functions listed in Table 1. PNSQPSO are tested for comparison with PPSO (parallel PSO), PQPSO (parallel QPSO). These functions are all minimization problems with minimum objective function values zeros. The fitness value is set as function value and the neighborhood of a particle is the whole population.

In order to investigate whether the PNSQPSO algorithms are well or not, as in [9], for functions f1 ~ f4, three different dimension sizes are tested. They are dimension sizes: 10, 20 and 30. The maximal number of generations is set as 1000,

Table 1 Expression of five tested benchmark functions

Test function	Function expression	Search	Initial range
Sphere	$f_1(X) = \sum_{i=1}^{n} x_i^2$	[−100, 100]	[50, 100]
Rosenbrock	$f_2(X) = \sum_{i=1}^{n-1} (100 \cdot (x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	[−30, 30]	[15, 30]
Rastrigin	$f_3(X) = \sum_{i=1}^{n} (x_i^2 - 10 \cdot \cos(2\pi x_i) + 10)$	[−5.12, 5.12]	[2.56, 5.12]
Griewank	$f_4(X) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[−600, 600]	[300, 600]
Schaffer's	$f_5(X) = 0.5 + \frac{(\sin(\sqrt{x^2+y^2})^2}{(1.0+0.001(x^2+y^2))^2}$	[−100, 100]	[30, 100]

1500, and 2000 corresponding to the dimensions 10, 20, and 30 for five functions, respectively. For the last function, the maximum generation is 2000 corresponding to the dimensions is 2. The population size is 80.

In PQPSO and PNSQPSO algorithm, Contraction-Expansion (CE) Coefficient decreases from 1.0 to 0.5 linearly.

4.1 Results

We had 50 trial runs for every instance and recorded mean best fitness and mean running time (unit: second).

The mean best fitness values and mean running time for 50 runs of each function are recorded in Tables 2, 3, 4, 5, and 6. From the comparison of PNSQPSO with other algorithms, PNSQPSO achieved better results than other algorithms both performance and efficiency. PNSQPSO has higher convergence precision and faster operation efficiency. PQPSO has achieved better results than PPSO algorithms averagely. In some case, PPSO has better performance and faster efficiency than PQPSO.

Figure 1 shows the convergence process of the three algorithms on the benchmark Rosenbrock and Rastrigin functions with swam size 80 and dimension 30 and CPU 8 averaged on 50 trail runs. It is shown that, although PNSQPSO converge more slowly than the PPSO and PQPSO during the early stage of search, it may catch up with PPSO and PQPSO at later stage and could be generated better solutions at the end of search.

Figure 2 illustrates the mean running time of the three algorithms on the functions f1 ~ f4 with swarm size 80 and dimension 30 averaged on 50 trail run. It is

Table 2 Simulation results of Sphere function

CPU	Dim	Gmax	PPSO		PQPSO		PNSQPSO	
			Fitness	Time	Fitness	Time	Fitness	Time
1	10	1000	1.60e-028	0.58	2.56e-103	0.41	2.73e-105	0.32
	20	1500	6.38e-018	2.09	1.41e-068	1.95	1.44e-069	1.42
	30	2000	3.27e-013	3.17	6.57e-050	3.06	3.85e-051	2.73
2	10	1000	7.92e-043	0.24	0	0.21	0	0.16
	20	1500	4.56e-034	1.96	3.14e-092	1.28	4.27e-134	1.05
	30	2000	2.94e-029	2.59	6.27e-073	2.03	7.83e-098	1.22
4	10	1000	6.51e-088	0.17	0	0.18	0	0.14
	20	1500	2.90e-052	1.07	3.24e-109	1.08	0	0.91
	30	2000	5.27e-034	1.13	5.37e-086	1.20	0	1.04
8	10	1000	5.34e-102	0.05	0	0.09	0	0.02
	20	1500	4.05e-082	0.76	0	0.83	0	0.03
	30	2000	8.77e-067	0.94	0	1.01	0	0.07

Table 3 Simulation results of Rosenbrock function

CPU	Dim	Gmax	PPSO		PQPSO		PNSQPSO	
			Fitness	Time	Fitness	Time	Fitness	Time
1	10	1000	19.0259	0.69	6.8312	0.54	3.2751	0.48
	20	1500	40.2289	2.15	33.5287	1.60	18.4602	1.37
	30	2000	56.8773	3.78	44.5946	3.33	30.4164	3.21
2	10	1000	10.7520	0.38	4.2516	0.27	2.9830	0.21
	20	1500	35.8567	1.16	20.4309	0.95	12.5482	0.87
	30	2000	42.5237	2.39	35.3771	2.16	25.7419	1.94
4	10	1000	8.4261	0.18	3.1067	0.21	2.1452	0.15
	20	1500	26.7024	0.52	10.0578	0.60	5.8937	0.36
	30	2000	33.5247	1.38	21.4814	1.47	12.3091	1.17
8	10	1000	4.3719	0.09	2.3061	0.15	1.1673	0.07
	20	1500	14.8529	0.28	8.5427	0.49	3.2871	0.24
	30	2000	21.5817	1.06	11.9046	1.17	8.0326	0.82

Table 4 Simulation results of Rastrigin function

CPU	Dim	Gmax	PPSO		PQPSO		PNSQPSO	
			Fitness	Time	Fitness	Time	Fitness	Time
1	10	1000	2.3890	0.52	2.2617	0.48	0.8630	0.33
	20	1500	12.8594	1.85	8.4121	1.60	4.3269	1.39
	30	2000	30.2140	2.82	14.8574	2.65	12.7531	2.14
2	10	1000	1.1132	0.20	1.0562	0.17	0.5436	0.13
	20	1500	10.9572	0.75	6.4308	0.64	2.8642	0.47
	30	2000	22.3041	1.85	10.4989	1.53	6.3719	1.26
4	10	1000	0.3857	0.11	0.1273	0.18	0.0821	0.11
	20	1500	3.8472	0.35	2.6418	0.47	1.4329	0.29
	30	2000	10.5618	0.96	8.2160	1.07	4.2160	0.85
8	10	1000	0.1506	0.08	0.0857	0.12	0.0011	0.07
	20	1500	2.8146	0.27	1.2573	0.31	0.2795	0.21
	30	2000	7.2381	0.50	5.2147	0.62	0.9846	0.38

shown that PNSQPSO algorithm has the fastest operation efficiency. The speedup ratio of PNSQPSO algorithm is much than PPSO and PQPSO algorithm. Parallel of PNSQPSO algorithm has better performance and efficiency. Along with the increase of the number of processors, the speedup ratio of PQPSO algorithm is less than PPSO algorithm.

Figure 3 shows the mean running time of PNSQPSO algorithms on the functions f1 and f5, f4 and f5 with swarm size 80 and dimension 30 and 2 averaged on 50 trail run. From comparison the comparison of f1 with f5, f4 with f5, we can draw the

Table 5 Simulation results of Griewank function

CPU	Dim	Gmax	PPSO		PQPSO		PNSQPSO	
			Fitness	Time	Fitness	Time	Fitness	Time
1	10	1000	0.0658	0.50	0.0052	0.41	0.0032	0.35
	20	1500	0.1034	1.19	0.0137	1.05	0.0086	0.92
	30	2000	0.1950	2.73	0.0371	2.18	0.0194	1.91
2	10	1000	0.0124	0.25	0.0029	0.20	0.0012	0.15
	20	1500	0.0852	0.48	0.0074	0.44	0.0057	0.36
	30	2000	0.1027	1.59	0.0108	1.36	0.0135	1.19
4	10	1000	0.0086	0.16	0.0010	0.18	0.0004	0.12
	20	1500	0.0248	0.40	0.0035	0.41	0.0012	0.23
	30	2000	0.0961	1.01	0.0062	1.15	0.0023	0.86
8	10	1000	0.0014	0.11	0.0002	0.12	0.0001	0.07
	20	1500	0.0082	0.29	0.0017	0.30	0.0003	0.18
	30	2000	0.0104	0.93	0.0024	0.98	0.0015	0.46

Table 6 Simulation results of Schaffer's function

CPU	Dim	Gmax	PPSO		PQPSO		PDCQPSO	
			Mean best	Time	Mean best	Time	Mean best	Time
1	2	2000	0	0.27	7.56e-009	0.21	7.82e-014	0.18
2	2	2000	0	0.18	4.35e-012	0.10	4.92e-031	0.06
4	2	2000	0	0.10	8.65e-025	0.08	0	0.04
8	2	2000	0	0.07	6.52e-056	0.04	0	0.01

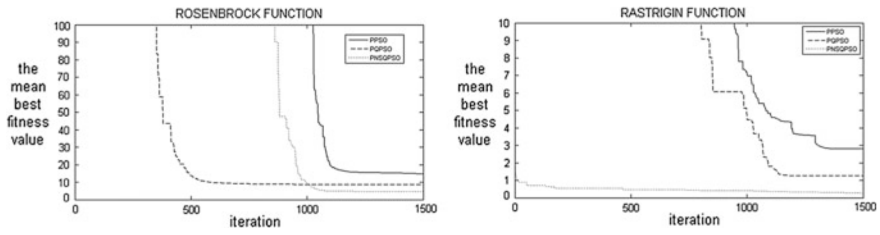


Fig. 1 Convergence process of the three algorithms on the Rosenbrock and Rastrigin functions

following conclusions; functions f1 and f4 decreased faster than function f5. When the problem complexity increases, the speedup ratio can be better.

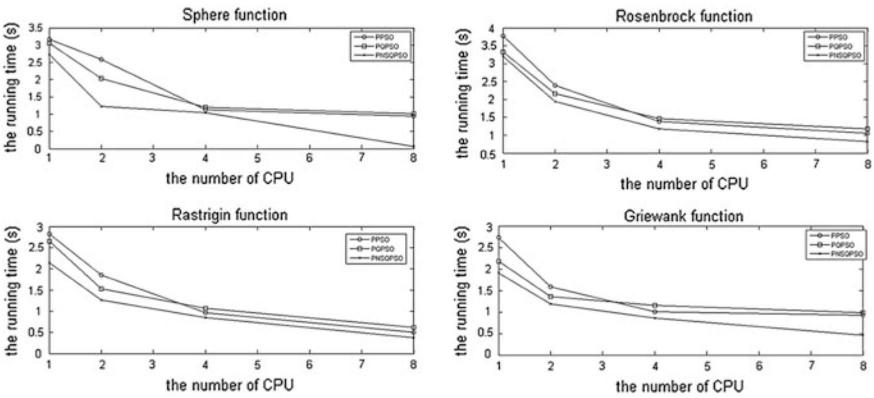


Fig. 2 The mean running time of the three algorithms on the functions f1 ~ f4 with swarm size 80 and dimension 30 averaged on 50 trail run

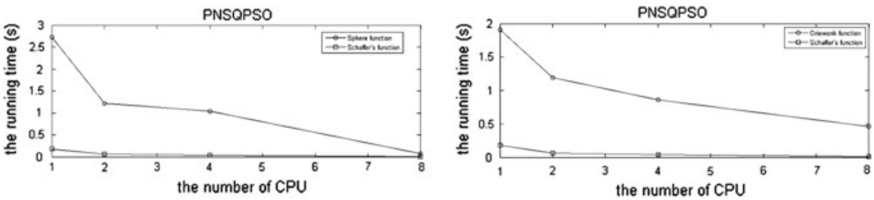


Fig. 3 The mean running time of PNSQPSO algorithms on the functions f1 and f5, f4 and f5 averaged on 50 trail run

5 Conclusions

This paper presents parallel o QPSO algorithm to solve complex optimization problems. The proposed approach explores diversity enhancing mechanism to increase diversity of swarm and avoid premature convergence of QPSO. To verify the performance of PNSQPSO, different types of benchmark functions are tested in the experiments, which show that PNSQPSO has better performance and higher efficiency than PPSO and PQPSO algorithm. Along with the increase of the number of processors, the speedup is more close to liner.

Acknowledgements Thanks to the support by the National Natural Science Fund (No. 71361008, No. 71461008), and Hainan Province key projects (No.ZDXM20130080), and the Hainan Province Natural Science Fund (No. 20151003).

References

1. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
2. Van den Bergh, F.: An Analysis of Particle Swarm Optimizers. University of Pretoria, South Africa (2001)
3. Sun, J., Wu, X. J., Palade, V., et al.: Convergence analysis and improvements of quantum-behaved particle swarm optimization. Inf. Sc. **193**, 81–103 (2012)
4. Lu, S.F., Sun, C.F., Lu, Z.D.: An improved quantum-behaved particle swarm optimization method for short-term combined economic emission hydrothermal scheduling. Energy Convers. Manage. **51**(3), 561–571 (2010)
5. Sun, J., Fang, W., Vasile, P.: Quantum-behaved particle swarm optimization with Gaussian distributed local attractor point. Appl. Math. Comput. **218**(7), 3763–3775 (2011)
6. Niu, Q., Zhou, Z., Zhang, H.Y., Deng, J.: An improved quantum-behaved particle swarm optimization method for economic dispatch problems with multiple fuel options and valve-points effects. Energies **5**(9), 3655–3673 (2012)
7. Sun, J., Chen, W., Fang, W., et al.: Gene expression data analysis with the clustering method based on an improved quantum-behaved particle swarm optimization. Eng. Appl. Artif. Intell. **25**(2), 376–391 (2012)
8. Wang, H., Sun, H., Li, C.H., et al.: Diversity enhanced particle swarm optimization with neighborhood search. Inf. Sci. **223**, 119–135 (2013)
9. Long, H.X., Sun, J., Wang, X.G., et al.: Using selection to improve quantum-behaved particle swarm optimization. Int. J. Innovative Comput. Appl. **2**(2), 100–114 (2010)

Intelligent Vehicle Scheduling with Soft Time Window and Non-full Load Constraint

Hai-xia Long, Cheng-yi Zhang and Hai-yan Fu

Abstract Vehicle scheduling belongs to a typical TSP (Traveling salesman problem) problem in the field of combinatorial optimization. We design a vehicle scheduling model with soft time window and non-full load constraints. The model belongs to single-objective optimization with multiple constrains. This article uses binary particle swarm optimization algorithm (BPSO) to solve the vehicle scheduling model, combining constraint deviation value method for solving constraint. Cross experiments are designed based on the number of customers and the number of delivery vehicles. The experiment results show that BPSO algorithm in short running time can resolve vehicle scheduling problem, and get a more satisfactory solution to objective function, the number of vehicles customers required, the distribution distance and path of the delivery vehicle. Thus it lays a technical foundation for intelligent vehicle scheduling.

Keywords Binary particle swarm optimization algorithm · Constraints deviation value · Vehicle scheduling · Delivery distance · Delivery path

1 Introduction

Vehicle Scheduling belongs to a typical TSP (Traveling salesman problem) problem in the field of combinatorial optimization, which is also the traveling salesman problem. This problem not only has too many constraints, but also is a NP problem with higher computational complexity. In recent decades, with the development of computer technology, some complex optimization problems that can not be solved in the past have obtained approximate solutions by swarm intelligence optimization

H. Long · C. Zhang (✉) · H. Fu
School of Information Science Technology, Hainan Normal University,
Haikou 571158, China
e-mail: chengyizh@hainnu.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_46

algorithm, so intelligent optimization algorithm for solving complex problems is becoming more and more important [1]. The swarm intelligence optimization algorithm including genetic algorithms (GA) and particle swarm optimization (PSO), due to their high versatility and no requirements for objective function properties, have become the main method to solve complex optimization problems.

Some scholars have conducted research for the public transport vehicle scheduling problem. Scholars in Brazil such as Maikol M. Rodriguesa, Cid C. Sottzab, and Arnaldo V. Moura [2], give the optimized vehicle plan to reduce operating costs under the passenger demand and constraints of technology conditions. Xu et al. established the vehicles scheduling optimization model with the object of the minimum vehicles quantity and the minimum total operating costs [3]. Homayouni et al. [4] presented a genetic algorithm (GA) to solve automated guided vehicles. Xiao et al. proposed a new mathematical formulation to describe the scheduling problem with hierarchical objectives and weighted tardiness [5].

Considering a variety of factors that affect vehicle scheduling problem, we establish an intelligent vehicle scheduling model with soft time windows and non-full load constraints, and use binary particle swarm optimization (BPSO) algorithm to solve the model. PSO is a kind of stochastic optimization algorithms proposed by Kennedy and Eberhart [6] that can be easily implemented and is computationally inexpensive. Several authors developed strategies to improve PSO. Deepa and Sugumaran improve the PSO evolutionary formula to effectively solve the problem of higher order linear equations [7]; Epitropakis et al. proposed PSO algorithm based on social and cognitive experience [8]; Mousa et al. proposed PSO algorithm based on local search strategy for multi-objective optimization problem [9]; Kennedy and Eberhart introduced discrete binary PSO (BPSO) [10], which can be applied to discrete binary variables.

2 Vehicle Scheduling Model with Soft Time Window and Non-full Load Constraints

2.1 Symbol Description

m is the number of delivery vehicles in distribution network;
F is the vehicle fix usage fee;
L is the maximum travel distance of the vehicle;
q_i is the demand of the ith customer, $s(q_i)$ is the minimum number of vehicles customer required

$$s(q_i) = \begin{cases} [q_i/Z_0], & \text{if } Z_0 \text{ is divisible by } q_i \\ [q_i/Z_0] + 1, & \text{if } Z_0 \text{ is not divisible by } q_i \end{cases} \tag{1}$$

n is the number of customer;
r_{ij} is distance from custom i to custom j;
c_{ij} represents the generalized cost from the customer i to j;
S is set of arbitrary customers;
V is set of all customers

2.2 Decision Variable

$$x_{ijk} = \begin{cases} 1, & \textit{if vehicle k is move from customer i towards customer j} \\ 0, & \textit{other} \end{cases}$$
$$y_{ik} = \begin{cases} 1, & \textit{if the tasks of customer i is performed by vehicle k} \\ 0, & \textit{other} \end{cases} \tag{2}$$

2.3 Vehicle Scheduling Model with Soft Time Window and Non-full Load Constraint

$$Min \quad T = Fm + \sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=1}^{m} c_{ij}x_{ijk} + \sum_{i=1}^{n} P(t_i), \tag{3}$$

ST:

$$\max \sum_{i=1}^{n} q_i y_{ik} \le Z_0, \; k = 1, 2 \ldots m \tag{4}$$

$$\sum_{i=0}^{n}\sum_{j=1}^{n} r_{ij}x_{ijk} \le L, \; k = 1, 2 \ldots m \tag{5}$$

$$\sum_{k=1}^{m} y_{ik} \le s(q_i), i = 1, 2 \ldots n, \tag{6}$$

$$\sum_{k=1}^{m} y_{0k} = m, \tag{7}$$

$$\mathrm{t}_0 = \mathrm{T}_0 = 0 \tag{8}$$

$$t_i > 0, \; i = 1, 2 \ldots n \tag{9}$$

$$T_i > 0, \ i = 1, 2 \ldots n \tag{10}$$

$$ET_i \le t_i \le LT_i, \ i = 1, 2 \ldots n \tag{11}$$

$$t_i + T_i + t_{ij} + (1 - x_{ijk})\mathrm{T} \le t_j, \ i = 1, 2 \ldots n, \ j = 1, 2 \ldots n, \ i \ne j \tag{12}$$

$$\sum_{i=1}^{n} x_{ijk} = y_{jk}, \ j = 1, 2 \ldots n, \ k = 1, 2 \ldots m, \tag{13}$$

$$\sum_{j=1}^{n} x_{ijk} = y_{ik}, \ i = 1, 2 \ldots n, \ k = 1, 2 \ldots m, \tag{14}$$

$$x_{ijk}(x_{ijk} - 1) = 0, \ i = 1, 2 \ldots n, \ j = 1, 2 \ldots n, \ k = 1, 2 \ldots m, \tag{15}$$

$$y_{ik}(y_{ik} - 1) = 0, \ i = 1, 2 \ldots n, \ k = 1, 2 \ldots m, \tag{16}$$

$$\sum_{i,j \in S} x_{ijk} \ge 1, \ \forall S \subseteq V, \ k = 1, 2 \ldots m, \tag{17}$$

where, $P(t_i)$ is penalty costs of the *i*th customer; t_i represents the time that distribution vehicle needs to reach the *i*th customer; T_i is the time that loading and unloading spent on the *i*th customer. t_{ij} is the time that vehicle needs to move from the *i*th customer to *j*th customer. Equation (3) shows the total of costs of distribution and penalties for violations the window are minimum; Eq. (4) indicates that the vehicle k undertaken distribution tasks by an amount no greater than the load of the vehicle; Eq. (5) indicates a delivery distance of the vehicle can not exceed the maximum travel distance of the vehicle; Eq. (6) represents the client i at most one vehicle is not loaded with its distribution services; Eq. (7) indicates that the amount m cars start from the distribution center; Eqs. (8)–(12) shows time window constraints; Eq. (8) represents the travel time and loading and unloading times of the vehicle in distribution center are 0; Eq. (9) represents the time that the vehicle reach the customer i is non-negative; Eq. (10) represents the time that the customer i for loading and unloading is non-negative; Eq. (11) indicating that the time the vehicle reaches the customer i should meet their time windows; Eq. (12) represents the time calculated formula from the customer i arriving to the customer j, that is the time relationship between precursors node and subsequent node; Eqs. (13) and (14) shows the relationship between two variables; Eqs. (15) and (16) represents the both variables are binary variables constraints; Eq. (17) indicating that circuits of delivery path are canceled.

3 BPSO Algorithm in Solving Vehicle Scheduling Model

3.1 BPSO Algorithm

In this paper, BPSO is chosen since the position of each particle can be given in binary string form (0 or 1), which adequately reflects the straightforward yes/no choice as to whether or not a feature needs to be selected. The changes in particle velocity and trajectory can be interpreted as a change in the probability of finding the particle in one state or another; since this change is a probability, it is limited to a range of $\{0.0 \sim 1.0\}$.

Based on the principles of PSO, we set the required particle number first, and then the initial coding string for each particle is generally created in such a way that the population of the particles is distributed randomly over the search space. In this study, we code each particle to imitate an SNP barcode in a BPSO. Each particle is coded to a binary string in which the bit value $\{1\}$ or $\{0\}$ represents a selected or non-selected feature respectively.

At every iteration, the particle's trajectory is updated by the binary evaluation of two "best" values, called *pbest* and *gbest*. The coordinates of each particle trajectory are associated with the best solution (fitness) the particle has achieved so far. And this fitness value is stored, and called *pbest*. When a particle takes the whole population as its topological neighbor, the best solution is a global "best" solution called *gbest*. Once the adaptive values *pbest* and *gbest* are obtained, the features of the *pbest* and *gbest* particles can be tracked with regard to their position and velocity. Each particle is updated according to the following equations.

$$v_{id}^{new} = w \times v_{id}^{old} + c_1 \times rand_1 \times (pbest_{id} - x_{id}^{old}) + c_2 \times rand_2 \times (gbest_i - x_{id}^{old}) \quad (18)$$

$$if \quad v_{id}^{new} \notin (V_{\min}, V_{\max}) \ then \quad v_{id}^{new} = \max(\min(V_{\max}, v_{id}^{new}), V_{\min}) \quad (19)$$

$$S(v_{id}^{new}) = \frac{1}{1 + e^{-v_{id}^{new}}} \quad (20)$$

$$if \quad (S(v_{id}^{new}) > rand) \ then \quad x_{id}^{new} = 1; \ else \quad x_{id}^{new} = 0 \quad (21)$$

where w is the inertia weight, c_1 and c_2 are acceleration (learning) factors, and *rand*, $rand_1$ and $rand_2$ are random numbers in the interval (0, 1). v_{id}^{new} and v_{id}^{old} are velocities for those updated particles and the particles before being updated, respectively, x_{id}^{old} is the original particle position (solution), and x_{id}^{new} is the updated particle position (solution). In Eq. (19), particle velocities of each dimension are tried to a maximum velocity $V_{\max}$. If the sum of accelerations causes the velocity of that dimension to exceed $V_{\max}$, then the velocity of that dimension is limited to $V_{\max}$. $V_{\max}$ and $V_{\min}$ are user-specified parameters (in our case $V_{\max} = 1$, $V_{\min} = 0$).

3.2 The Method of Constrained Deviation Value

Typically, processing constraints have two ways: the constraint deviation value and the constraint deviation degree. In this paper, we use the constraint deviation value [11].

Standard constraint deviation value is calculated as follows:

$$C(x) = \sum_{j=1}^{q} w_j \frac{c_j(x)}{c_j^{\max}} \tag{22}$$

where $c_j(x)$ is the jth constraint deviation value, which is calculated as follows in Eq. (23); $c_j^{\max}$ is the maximum of jth constraint deviation value,which is defined in Eq. (24); $w_j = \frac{1}{q}$; $C(x)$ is the constraint deviation value of individual x.

$$c_j(x) = \begin{cases} \max(0, g_j(x)), & j = 1, \ldots, q \\ \max(0, |h_j(x)| - \varepsilon), & j = q+1, \ldots, m \end{cases} \tag{23}$$

$$c_j^{\max} = \max_x c_j(x) \tag{24}$$

Processing constraints deviation value is divided into two types: (1) the constraint deviation value is added to each objective function value of each individual, which simplifies the constraint problem to an unconstrained problem; (2) regarding the constraint deviation value as a dimensional objective function value, so that the objective function becomes k + 1 dimensions after the new objective function can be optimized in non-constraint.

4 Experiments and Results

4.1 Initialization

The vehicle has a fixed cost of 200 yuan, the maximum travel distance of the vehicle is 50 km, and customer demand is located between [50,400] 100 kg. The decision variables x_{ijk} and y_{ik} are initialized based on the number of customer and

Fig. 1 The value of y_{ik}

```
y_ik =
```

1	1	1	1	0	0	0
1	0	0	0	1	0	1
0	0	1	1	0	1	1
1	0	0	1	0	0	1
0	0	1	0	1	1	0
0	0	0	1	1	1	0
0	1	0	1	0	0	0
0	1	1	1	0	1	1
0	1	1	1	0	0	0
0	0	0	1	0	0	1
0	0	1	0	1	1	0
0	0	1	0	0	1	1
1	1	0	0	1	1	0
1	0	0	0	1	0	1
0	0	0	1	0	0	1
0	0	0	0	0	1	1
0	1	1	0	0	1	1
1	1	1	0	0	0	0
0	1	1	0	0	0	1
0	0	1	0	1	1	0

the delivery vehicles. If the number of customer is 20 and the number of delivery vehicles is 7, the decision variables x_{ijk} are seven matrix with size of 20×20, the decision variables y_{ik} is a matrix with size of 20×7. Figure 1 shows the value of y_{ik}. Time for loading and unloading cargo is in interval [60, 180].

In our experiment, the BPSO termination condition reaches at a pre-specified number of iterations (in our case the number of iterations was 1000). Population size is 20, the inertia weight w is decreasing linearly from 0.9 to 0.4, and acceleration (learning) factors C_1 and C_2 are set to 2.

4.2 Results

The number of customer is 30, the demand of each customer is 363, 143, 243, 64, 107, 148, 309, 98, 113, 361, 339, 350, 369, 198, 270, 268, 190, 170, 235, 141, 308, 380, 357,130, 398, 70, 398, 56, 71, 255 respectively. The number of distribution vehicles is 4, 5, 6, 7 respectively. Table 1 summarizes the objective function value. Figure 2 illustrates the convergence of BPSO algorithm.

A delivery distance of delivery vehicle is shown in Table 2, and the bar chart is shown in Fig. 3.

Table 1 Objective function value (unit: yuan)

	Vehicle = 4	Vehicle = 5	Vehicle = 6	Vehicle = 7
Fitness value	3202	3362	3511	3803 元

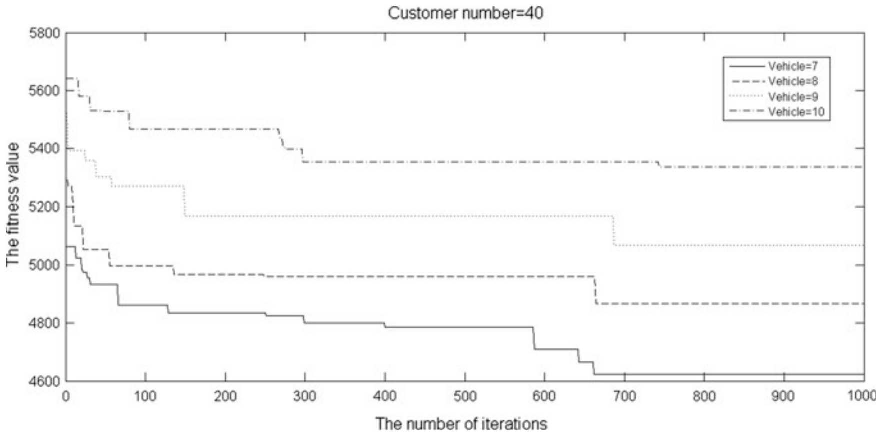


Fig. 2 The convergence of BPSO algorithm

Table 2 A delivery distance of delivery vehicle

	Distance
Vehicle = 4	46 46 34 50
Vehicle = 5	30 30 38 44 45
Vehicle = 6	26 31 20 20 36 43
Vehicle = 7	29 22 37 29 22 13 34

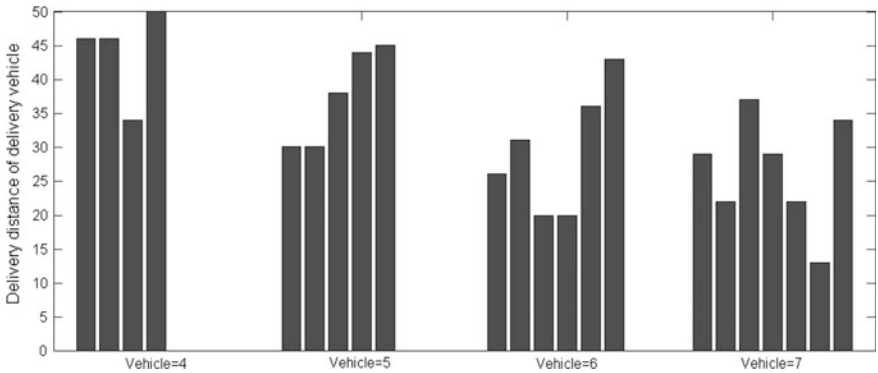


Fig. 3 A delivery distance of delivery vehicle

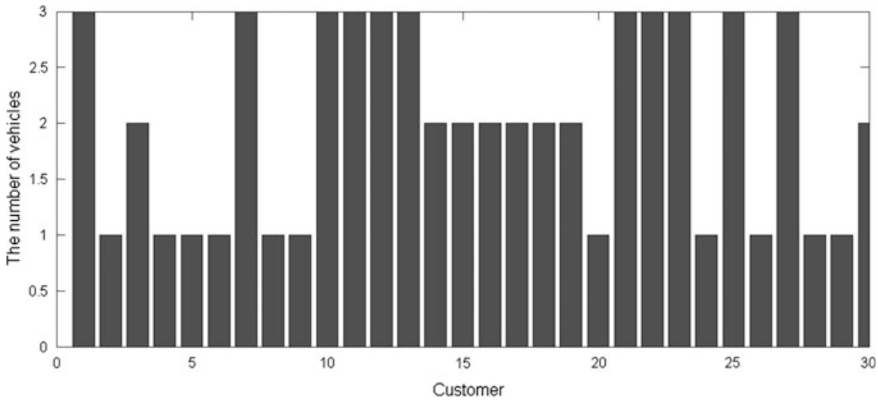


Fig. 4 Customer demand for the number of vehicles

The number of vehicles to customer needs is shown in Fig. 4.
Delivery path of delivery vehicle is described in Table 3.
The run time of BPSO algorithm is shown in Table 4.

Table 3 Delivery path of delivery vehicle

		Path
Vehicle = 4	The first vehicle	4 → 7→10 → 11 → 12 → 13 → 16 → 18 → 20 → 21 → 23 → 25 → 27 → 28 → 29
	The second vehicle	1 → 2→3 → 6→7 → 10 → 12 → 13 → 14 → 15 → 17 → 18 → 22 → 25 → 27
	The third vehicle	1 → 3→7 → 10 → 11 → 17 → 19 → 21 → 22 → 23 → 26 → 30
	The fourth vehicle	1 → 5→8 → 9→11 → 12 → 13 → 14 → 15 → 16 → 19 → 21 → 22 → 23 → 24 → 25 → 27 → 30
Vehicle = 5	The first vehicle	1 → 3→6 → 7→9 → 11 → 14 → 15 → 19 → 21 → 24 → 25
	The second vehicle	5 → 8→11 → 12 → 13 → 17 → 18 → 20 → 21 → 22 → 30
	The third vehicle	1 → 3→7 → 10 → 11 → 15 → 16 → 18 → 23 → 26 → 27
	The fourth vehicle	1 → 4→10 → 12 → 13 → 14 → 16 → 19 → 22 → 23 → 25 → 27
	The fifth vehicle	2 → 7→10 → 12 → 13 → 17 → 21 → 22 → 23 → 25 → 27 → 28 → 29 → 30
Vehicle = 6	The first vehicle	1 → 3→6 → 10 → 16 → 19 → 21 → 24 → 25 → 27 → 30
	The second vehicle	1 → 8→11 → 12 → 13 → 14 → 15 → 17 → 18 → 22 → 23
	The third vehicle	4 → 9→11 → 12 → 20 → 21 → 23
	The fourth vehicle	5 → 7→13 → 14 → 16 → 22 → 26
	The fifth vehicle	3 → 7→10 → 11 → 15 → 19 → 22 → 25 → 27 → 28 → 29 → 30
	The sixth vehicle	1 → 2→7 → 10 → 12 → 13 → 17 → 18 → 21 → 23 → 25 → 27
Vehicle = 7	The first vehicle	1 → 12 → 13 → 14 → 18 → 23 → 25 → 30
	The second vehicle	10 → 14 → 15 → 17 → 20 → 21 → 22 → 23
	The third vehicle	1 → 3→5 → 11 → 15 → 16 → 21 → 25 → 27 → 28
	The fourth vehicle	1 → 3→6 → 7→8 → 11 → 17 → 18 → 22 → 23 → 26
	The fifth vehicle	12 → 13 → 19 → 22 → 24 → 25 → 30
	The sixth vehicle	2 → 7→10 → 11 → 16 → 27
	The seventh vehicle	4 → 7→9 → 10 → 12 → 13 → 19 → 21 → 27 → 29

Table 4 The run time of BPSO algorithm (unit: s)

Custom = 30			
Vehicle = 4	Vehicle = 5	Vehicle = 6	Vehicle = 7
79.9044	89.8471	105.7972	114.8382

5 Conclusions

In this paper, we introduce an approach of intelligent vehicle scheduling with binary particle swarm optimization and constraints deviation value. The results of our experiments show that the BPSO is a remarkably effective method for vehicle scheduling.

From the above results shown in tables and figures, we can obtain the following conclusion: when the number of customer is known, the more distribution vehicle it has, the higher cost it spends; the more vehicles start from the distribution center, the shorter delivery distance the vehicle has, and a delivery distance of delivery vehicle does not exceed the maximum travel distance of 50 km; if the number of customer is known, the fewer delivery vehicles, the longer the distribution route; if the number of distribution vehicles is known, the more customer number, the higher distribution cost.

From the running time of the BPSO algorithm, we state that with the increase of the customer and the distribution vehicle, the running time of the algorithm increases linearly. When the customer number is 20, the average running time of the BPSO algorithm is 69.1353 s (about 1 min), and when the number of customers is 50, the average running time of the BPSO algorithm is 242.6622 s (about 4 min). Therefore, BPSO algorithm can be used to solve the vehicle scheduling problem.

Acknowledgements This work was supported by the National Natural Science Fund (No.71361008, No. 71461008), and Hainan Province key projects (No.ZDXM20130080), and the Hainan Province Natural Science Fund (No. 20151003).

References

1. Ling, W.: Intelligent Optimization Algorithm and its Application [M]. Tsinghua University Press, Springer-Verlag, Beijing (2001)
2. Kliewer, N., Mellouli, T., et al.: A time space network based exact optimization model for multidepot bus scheduling. Eur. J. Oper. Res. **175**, 1616–1627 (2006)
3. Xu, Z.C., He, P.J., Teng, J., Liu, L.P.: Transit vehicles intelligent scheduling optimization based on the division of characteristic periods. Soc. Behav. Sci. **96**, 1502–1512 (2013)
4. Homayouni, S.M., Tang, S.H., Motlagh, O.: A genetic algorithm for optimization of integrated scheduling of cranes, vehicles, and storage platforms at automated container terminals. J. Comput. Appl. Math. **270**, 545–556 (2014)

5. Xiao, Y.Y., Konak, A.: A simulating annealing algorithm to solve the green vehicle routing and scheduling problem with hierarchical objectives and weighted tardines. Appl. Soft Comput. **34**, 372–388 (2015)
6. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, IV, pp. 1942–1948. IEEE Service Center, Piscataway, NJ (1995)
7. Deepa, S.N., Sugumaran, G.: Model order formulation of a multivariable discrete system using a modified particle swarm optimization approach. Swarm Evol. Comput. **1**(4), 204–212 (2011)
8. Epitropakis, M.G., Plagianakos, V.P., Vrahatis, M.N.: Evolving cognitive and social experience in particle swarm optimization through differential evolution: a hybrid approach. Inf. Sci. **216**, 50–92 (2012)
9. Mousa, A.A., El-Shorbagy, M.A., Abd-El-Wahed, W.F.: Local search based hybrid particle swarm optimization algorithm for multiobjective optimization. Swarm Evol. Comput. **3**, 1–14 (2012)
10. Kennedy, J., Eberhart, R.C.: A discrete version of the particle swarm algorithm. In: Proceedings of the 1997 Conference on System, Man and Cybernetics, pp. 4104–4109. IEEE Service Center, Piscataway, NJ (1997)
11. Venkatraman, S., Yen, G.G.: A generic framework for constrained optimization using genetic algorithms. IEEE Trans. Evol. Comput. **9**(4), 424–435 (2005)

Two-Stage Flow Distribution Model Based on Multi-objective Optimization

Cheng-yi Zhang, Hai-xia Long and Hai-yan Fu

Abstract Efficiency and effectiveness of the logistics network is directly dependent on network topology and flow distribution of logistics nodes and transport routes. Flow distribution model based on multi-objective function belongs to NP- difficult problem. The model consists of supply, demand and constraints of network capacity and so on. The number of network paths will exponentially expand with the size of the network, that is, when the number of various types of nodes increases, the overall transportation route will rapidly increase. So we design a "sliced, two-phase" flow distribution model of logistics network. Combining with penalty-parameter-less method, we use particle swarm optimization (PSO) algorithm to optimize the model so as to find the solution of multi-objective function. Experiment results show that PSO algorithm with penalty function method can obtain satisfactory solution of flow distribution and network usage in a short period of time.

Keywords Logistics network · Flow distribution · Multi-objective function · Penalty function

1 Introduction

In practice, each logistics nodes and transport routes of the logistics network are concerned with a number of evaluation criteria. During the flow distribution of logistics network, we should consider the delivery time of goods, transport distance, the network reliability and so on, apart from the logistics cost. Thus, the flow distribution of the logistics network is actually the problem of multi-objective decision making, namely, to find multiple targets overall optimal flow distribution scheme.

C. Zhang · H. Long (✉) · H. Fu
School of Information Science Technology, Hainan Normal University,
Haikou 571158, China
e-mail: haixia_long@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_47

With the introduction of multi-objective and multi-objective optimization algorithm, relevant researches of logistics network optimization are made from single-objective optimization to multi-objective optimization [1–5]. Multi-objective usually considers cost, profit, distance, time, timeliness and reliability. Reference [6] discusses logistics network allocation of two goals such as cost and distance, and proposes model conversion method similar to lexicographic ordering algorithm. Reference [7] studies fuzzy transport problems of multi-objective and multi-mode, and designs fuzzy compromise programming method for solving this problem. Reference [8] studies center location of flow distribution problem with logistics cost and time reliability as objective, and converts the multi-objective to single-objective based on constraint method, and then uses greedy heuristic algorithm to solve problem. Reference [9] studies the inventory routing problem with time windows under conditions of random demand, establishes inventory and distribution strategy optimization model with inventory costs, order goods costs and distribution transport costs as goals, and uses genetic algorithm with elitist to solve the model. Reference [10] considers the needs and distribution path uncertain, and studies the emergency logistics and material distribution facility location. Reference [11] studies fuzzy multi-objective location—routing problem of the emergency logistics, establishes a model with goods delivery time and cost as goals, and uses genetic algorithms to solve problems.

Logistics network flow models established in the above references have drawbacks, that is, the number of the total transport routes will rapidly increase with the number of various types of nodes increases. So we design a "sliced, two-stage" logistics network flow distribution model. We combine particle swarm optimization (PSO) with the first method to solve traffic model.

2 "Sliced, Two-Phase" Flow Distribution Model of Logistics Network

2.1 *Model Hypothesis*

(1) Goods transported from the supply node to required node must via intermediate node. Intermediate node has part of storage capacity in addition to cargo intermediate and distribution functions;
(2) Between each, supply node, intermediate node and demand node, there must have only one transport route to connect them, without direct transport routes between intermediate nodes;
(3) Demand for goods of the demand node is known;
(4) Supply cargo capacity of supply node and transport capacity of transport route are determined;
(5) Cargo via the transport routes and the time consumed in logistics nodes are determined.
(6) Freight vehicles have the same load Z_0.

2.2 Symbol Description

(1) Upper and lower standard
s_i indicates supply node, i = 1, 2,…, I; d_j indicates demand node, j = i, 2,…, J; m_k indicates intermediate node, k = 1, 2,…, K; L_{ik}^{sm} indicates transport routes from supply node s_i to intermediate node m_k, i = l, 2,…, I, k = l, 2,…, K; L_{kj}^{md} indicates transport routes from intermediate node m_k to demand node d_j, k = l, 2,…, K, J = 1, 2,…, J.
(2) Costs
c_{ik}^{sm} indicates transportation cost from supply node s_i to intermediate node m_k; c_{kj}^{md} indicates transportation cost from intermediate node m_k to demand node d_j; c_k^m indicates intermediate cost via intermediate node m_k; c_k^h indicates inventory cost via intermediate node m_k; C represents the total cost of the logistics network of cargo;
(3) Time
t_{ik}^{sm} indicates transport time of cargo from supply node s_i to intermediate node m_k; t_{kj}^{md} indicates transport time of cargo from intermediate node m_k to demand node d_j; t_k^m indicates unit intermediate time that the goods in intermediate node u need; $t_{kj} = t_k^m + t_{kj}^{md}$ represents the time of the lower logistics network from intermediate node to demand node; $t_{ikj} = t_{ik}^{sm} + t_k^m + t_{kj}^{md}$ represents the time form supply node s_i to demand node d_j via intermediate node m_k; T represents one path delivery time of logistics network.
(4) Capacity
y_i^s indicates cargo supply capacity of supply node s_i; y_j^d indicates demand for cargo of demand node d_j;
y_k^m indicates circulation capacity of intermediate node m_k; y_k^h indicates storage capacity of intermediate node m_k; y_{ik}^{sm} indicates capacity of transport route L_{ik}^{sm}; y_{kj}^{md} indicates capacity of transport route L_{kj}^{md};
(5) Usage
u_{ik}^{sm} indicates usage of transport route L_{ik}^{sm}; u_{kj}^{md} indicates usage of transport route L_{kj}^{md}; u_k^m indicates usage of intermediate node m_k; u_k^h indicates an intermediate node existing inventory usage; u_i^s indicates usage of supply node s_i; U represents the total usage of the logistics network.
(6) Decision variables
x_{ik}^{sm} indicates logistics scale from supply node s_i to intermediate node m_k; x_{kj}^{md} indicates logistics scale from intermediate node m_k to demand node d_j; x_k^m indicates logistics scale via intermediate node m_k; x_k^h represents the amount of goods supplied directly to the demand node from the intermediate node.

2.3 Model Building

Considering capacity constraints in transport route and logistics node, we build following multi-objective function: logistics costs are minimized from supply node to demand node via intermediate node; transport time of the longest one way is minimum from supply node to demand node via intermediate node; and usage of logistics network is maximum.

The flow distribution model is shown in Eqs. (1)–(4).

$$\begin{aligned}\mathrm{Min\,C} = \sum_{i=1}^{I}\sum_{k=1}^{K} \mathrm{c}_{ik}^{sm}\mathrm{g}_{ik}^{sm}(\mathrm{x}_{ik}^{sm}) + \sum_{k=1}^{K}\sum_{i=1}^{I} \mathrm{c}_{kj}^{md}\mathrm{g}_{kj}^{md}(\mathrm{x}_{kj}^{md}) \\ + \sum_{k=1}^{K}(\mathrm{c}_{k}^{m}\mathrm{g}_{k}^{m}(\mathrm{x}_{k}^{m}) + \mathrm{c}_{k}^{h}\mathrm{g}_{k}^{h}(\mathrm{x}_{k}^{h}))\end{aligned} \tag{1}$$

$$\begin{aligned}\mathrm{Min\,T} = \max\{t_{ikj}: x_{ik}^{sm}x_{k}^{m}x_{kj}^{md} > 0\}, \mathrm{i} = 1, 2, \ldots, \mathrm{I}; \\ \mathrm{k} = 1, 2, \ldots, \mathrm{K}; \mathrm{j} = 1, 2, \ldots, \mathrm{J}.\end{aligned} \tag{2}$$

$$\mathrm{Max\,U} = \sum_{i=1}^{I}\sum_{k=1}^{K} \mathrm{u}_{ik}^{sm} + \sum_{k=1}^{K}\sum_{i=1}^{I} \mathrm{u}_{kj}^{md} + \sum_{i=1}^{I} \mathrm{u}_{i}^{s} + \sum_{k=1}^{K} \mathrm{u}_{k}^{m} + \sum_{k=1}^{K} \mathrm{u}_{k}^{h} \tag{3}$$

$$\begin{cases} \sum_{k=1}^{K} \mathrm{x}_{ik}^{sm} \le y_{i}^{s}, & i = 1, 2, \ldots, I \\ \sum_{k=1}^{K} \mathrm{x}_{kj}^{md} = y_{j}^{d}, & j = 1, 2, \ldots, J \\ x_{k}^{m} = \sum_{j=1}^{J} \mathrm{x}_{kj}^{md} = \sum_{k=1}^{K} \mathrm{x}_{kj}^{md} + x_{k}^{h}, & k = 1, 2, \ldots, K \\ x_{k}^{m} \le \sum_{k=1}^{K} \mathrm{x}_{kj}^{md} + y_{k}^{h} \le y_{k}^{m}, & k = 1, 2, \ldots, K \\ x_{ik}^{sm} \le y_{ik}^{sm} & i = 1, 2, \ldots, I, k = 1, 2, \ldots, K \\ x_{kj}^{md} \le y_{kj}^{md} & k = 1, 2, \ldots, K, j = 1, 2, \ldots, J \\ x_{ik}^{sm}, x_{kj}^{\mathrm{md}}, x_{k}^{m}, x_{k}^{h} \in N^{*}, & i = 1, 2, \ldots, I, k = 1, 2, \ldots, K, j = 1, 2, \ldots, J \end{cases} \tag{4}$$

N^{*} is non-negative integers.

Equation (1) indicates that logistics costs is minimized. Where $\sum_{i=1}^{I}\sum_{k=1}^{K} \mathrm{c}_{ik}^{sm}\mathrm{g}_{ik}^{sm}(\mathrm{x}_{ik}^{sm})$ indicates the transport costs form supply node to intermediate node; $\sum_{k=1}^{K}\sum_{i=1}^{I} \mathrm{c}_{kj}^{md}\mathrm{g}_{kj}^{md}(\mathrm{x}_{kj}^{md})$ indicates the transport costs from intermediate node to demand node; $\sum_{k=1}^{K}(\mathrm{c}_{k}^{m}\mathrm{g}_{k}^{m}(\mathrm{x}_{k}^{m}) + \mathrm{c}_{k}^{h}\mathrm{g}_{k}^{h}(\mathrm{x}_{k}^{h}))$ indicates that transfer costs and storage costs of cargo are intermediate node. In Costs function cg(x), c is costs factor, g = g(x) is a function of flow x. Taking into account the non-full load will result increasing of transportation costs, we set: $g(x) = C_0^* Z_0^* s(x)$, where

$$s(x) = \begin{cases} \left[\frac{x}{Z_0}\right], & \text{if } Z_0 \text{ is divisible by x} \\ \left[\frac{x}{Z_0}\right] + 1, & \text{if } Z_0 \text{ is not divisible by x} \end{cases} \quad (5)$$

is the minimum number required for the vehicle, C_0 is unit transportation costs, Z_0 is standard load.

Equation (2) indicates that the maximum one-way transit time is minimized. One-way transit time consists of three parts, namely transport time of goods from supply node to intermediate node, transit time in intermediate node, transport time of goods from intermediate node to demand node.

Equation (3) represents maximum usage of logistics network, that is the total sum of usage between supply node, intermediate node and demand node is maximize, in which calculation formula of usage is shown in Eq. (6).

$$u = \begin{cases} x/y, & x > 0 \\ 1 & x = 0 \end{cases} \quad (6)$$

Equation (4) is showing constraints, respectively supply node capacity constraints, the demand for demand node constraints, flow balance constraints of intermediate node, capacity constraints of intermediate node, capacity constraint of transport routes, as well as non-negative integer constraints of flow.

Assumptions: a logistics network

(1) The number of supply node is I. The supply ability of the supply node i is y_i^s (i = 1, 2,…, I); The total supply vector is: $ys = (y_1^s, y_2^s, \ldots, y_I^s)$;
(2) The number of intermediate node is K.
There are I * K transport routes from supply nodes to intermediate nodes.
(3) The number of demand node is J; The demand of the demand node j is y_j^d (j = 1, 2,…, J); The total demand vector is: $yd = (y_1^d, y_2^d, \ldots, y_J^d)$;

There are K * J transport routes from intermediate nodes to demand nodes. So, there are I * K * J transport routes from supply nodes to demand node via intermediate nodes. In practice, when the number of various types of nodes increases, the total number of transport routes rapidly increases (for example, I = 20, K = 4, J = at 100, I * K * J = 8000), which causes the massive increase of the flow model computation. To this end, we have designed a "sliced, two-stage" optimization model to solve this problem.

2.4 "Sliced" Method

According to the actual scheduling of fresh food, we can assume that the intermediate nodes and demand nodes in the suburbs are relatively evenly distributed, so we can distribute the demand nodes evenly to the near intermediate nodes in accordance with the "principle of proximity" and the intermediate nodes can distribute the goods to districts. Each intermediate node m_k needs to ensure that the service node is n_k,

That is: $D = D_1 \cup D_2 \cup \cdots \cup D_K$

$$D_1 = (d_{11}, d_{11}, \ldots, d_{1n_1}), D_2 = (d_{21}, d_{21}, \ldots, d_{2n_2}), \ldots,$$
$$D_K = (d_{K1}, d_{K2}, \ldots, d_{Kn_K})$$

The total demand is: $yD = yD_1 \cup yD_2 \cup \cdots \cup yD_K$ where $yD_1 = \left(y_1^{D_1}, y_2^{D_1}, \ldots, y_{n_1}^{D_1}\right)$,

$$yD_2 = \left(y_1^{D_2}, y_2^{D_2}, \ldots, y_{n_2}^{D_2}\right), \ldots, yD_K = (y_1^{D_K}, y_2^{D_K}, \ldots, y_{n_K}^{D_K})$$

$n_1 + n_2 + \cdots + n_K = J$; There are $n_1, n_2, \ldots, n_k$ transport routes respectively from each intermediate node to demand node. The total transport route is $n_1 + n_2 + \cdots + n_K = J$. The demand for demand node via intermediate node k for service is: $yD_k = (y_1^{D_k}, y_2^{D_k}, \ldots, y_{n_k}^{D_k})$, (k = 1, 2,…, K).

So, there are I * J transport routes from supply nodes to demand node via intermediate nodes.

In particular, when I = 20, K = 4, J = 100, if $n_1 = n_2 = n_3 = n_4 = 25$, then the number of transport route is 2000.

According to our distribution agreement, we decompose flow distribution of lower logistics network into K sub-distribution. The allocation model is as follows: $c_{kj}^{md_i}$ indicates transportation cost from intermediate node m_k to demand node $d_{ij}(j = 1, 2\ldots, n_i)$; $x_{kj}^{md_i}$ indicates the logistics scale from intermediate node m_k to demand node $d_{ij}(j = 1, 2,\ldots, n_i)$.

2.5 "Two-Stage" Method

We use two-stage solution to the flow distribution illustrated in Figs. 1 and 2, avoiding computational complexity excessive explosion with the number of paths increasing.

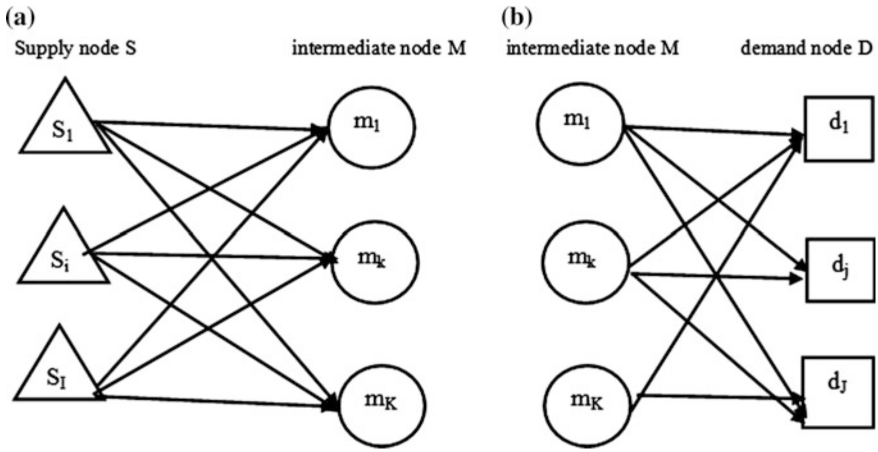


Fig. 1 Logistics network decomposition. **a** Upper logistics network, **b** lower logistics network

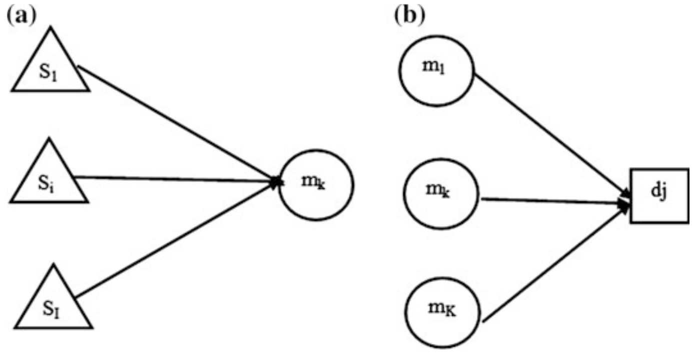


Fig. 2 Levels of logistics network decomposition. **a** Upper sub network, **b** lower sub network

(1) Lower logistics network flow distribution model

$$\text{Min}\, C_{low} = \sum_{p=1}^{K} \sum_{j=1}^{n_p} c_{pj}^{md_p} g_{pj}^{md_p} (x_{pj}^{md_p}) + \sum_{p=1}^{K} c_p^m g_p^m (x_p^m) \tag{7}$$

$$\text{Min}\, T = \max\{t_{pj}: x_{pj}^{md_p} > 0\},\ p = 1, 2, \ldots,, K; j = 1, 2, \ldots, n_p. \tag{8}$$

$$\text{Max}\, U = \sum_{p=1}^{K} \sum_{j=1}^{n_p} u_{pj}^{md_p} + \sum_{p=1}^{K} u_p^m \tag{9}$$

$$\text{s.t.}\begin{cases} x_{pj}^{md_p} = y_j^{d_p}, & j = 1, 2, \ldots, n_p, p = 1, 2, \ldots, K \\ x_p^m = \sum_{j=1}^{n_p} x_{pj}^{md_p}, & j = 1, 2, \ldots, n_p, p = 1, 2, \ldots, K \\ x_p^m \leq \sum_{i=1}^{I} x_{ip}^{sm} + y_p^h \leq y_P^m, & p = 1, 2, \ldots, K \\ x_{pj}^{md_p} \leq y_{pj}^{md_p} & j = 1, 2, \ldots, n_p, p = 1, 2, \ldots, K \\ x_{pj}^{md_p}, x_p^m \in N^*, & j = 1, 2, \ldots, n_p, p = 1, 2, \ldots, K \end{cases} \tag{10}$$

N^* is non-negative integers

(2) Upper logistics network flow distribution model

$$\text{Min } C_{up} = \sum_{i=1}^{I} \sum_{p=1}^{K} c_{ip}^{sm} g_{ip}^{sm}\left(x_{ip}^{sm}\right) + \sum_{p=1}^{K} c_p^h g_p^h(x_p^h) \tag{11}$$

$$\text{Min T} = \max\{t_{ip}: x_{ip}^{sm} > 0\}, \ \text{i} = 1, 2, \ldots, \text{I}; \ \text{p} = 1, 2, \ldots,, \text{K}. \tag{12}$$

$$\text{Max U} = \sum_{i=1}^{I} \sum_{p=1}^{K} u_{ip}^{sm} + \sum_{i=1}^{I} u_i^s + \sum_{p=1}^{K} u_p^h \tag{13}$$

$$\text{s.t.}\begin{cases} \sum_{p=1}^{K} x_{ip}^{sm} \leq y_i^s, & i = 1, 2, \ldots, I \\ x_p^m = \sum_{i=1}^{I} x_{ip}^{sm} + x_p^h, & p = 1, 2, \ldots, K \\ x_p^m \leq \sum_{i=1}^{I} x_{ip}^{sm} + y_p^h \leq y_p^m, & p = 1, 2, \ldots, K \\ x_{ip}^{sm} \leq y_{ip}^{sm} & i = 1, 2, \ldots, I, p = 1, 2, \ldots, K \\ x_{ip}^{sm} \in N^*, & i = 1, 2, \ldots, I, p = 1, 2, \ldots, K \end{cases} \tag{14}$$

N^* is non-negative integers

3 Model Solution Based on PSO and Penalty-Parameter-Less

3.1 PSO Algorithm

In the original PSO with M individuals [12], each individual is treated as an infinitesimal particle in the *D*-dimensional space, with the position vector and velocity vector of particle *i*, $X_i(t) = (X_{i1}(t), X_{i2}(t), \ldots, X_{iD}(t))$ and $V_i(t) = (V_{i1}(t), V_{i2}(t), \ldots, V_{iD}(t))$. The particle moves according to the following equations:

$$V_{ij}(t+1) = V_{ij}(t) + c_1 \cdot r_1 \cdot (P_{ij}(t) - X_{ij}(t)) + c_2 \cdot r_2 \cdot (P_{gj}(t) - X_{ij}(t)) \tag{15}$$

$$X_{ij}(t+1) = X_{ij}(t) + V_{ij}(t+1) \tag{16}$$

for $i = 1, 2, \ldots, M; j = 1, 2, \ldots, D$. The parameters c_1 and c_2 are called the acceleration coefficients. Vector $P_i = (P_{i1}, P_{i2}, \ldots, P_{iD})$ known as the *personal best position*, is the best previous position (the position giving the best fitness value so far) of particle i; vector $P_g = (P_{g1}, P_{g2}, \ldots, P_{gD})$ is the position of the best particle among all the particles and is known as the *global best position*. The parameters r_1 and r_2 are two random numbers distributed uniformly in (0, 1), that is r_1, $r_2 \sim U(0, 1)$. Generally, the value of V_{ij} is restricted in the interval $[-V_{\max}, V_{\max}]$.

3.2 Target Weighting Method

Flow distribution model is multi-objective function. In this paper we convert multi-objective optimization function into single-objective optimization based on target weighting method. The basic idea is to give each objective function for each weighting factor, then all the target components are multiplied by weighting factor, at last all components are coupled together to form a new single objective function. So we can use single objective optimization algorithm to find the solution of the model.

Target weighting formula is as follows:

$$F(x) = \sum_{i=1}^{m} \lambda_i f_i(x) \tag{17}$$

where $\lambda_i (i = 1, \ldots, m)$ is weighting factor of objective function $f_i(x)$, and $\sum_{i=1}^{m} \lambda_i = 1$.

3.3 Penalty-Parameter-Less for Constraint

This paper uses penalty-parameter-less [13] to handle constraint of flow distribution model. Each constraint has a penalty parameter, when the individual of population does not meet the constraints, the fitness function will be punished.

Multi-objective optimization problem is described as follows:

$$\begin{cases} Min\ y = F(x) = (f_1(x), f_2(x), \ldots, f_m(x)) \\ s.t. \quad g_i(x) \le 0, \quad i = 1, 2, \ldots, h; \\ h_i(x) = 0, \qquad i = h+1, h+2, \ldots, q \\ x \in X \in R^n, y \in Y \in R^m \end{cases} \tag{18}$$

For inequality constraints, we use penalty parameter $\max\{g_i(x), 0\}$; For equality constraints, we use penalty parameter $|h_i(x)|$.

4 Experiments and Results

4.1 Experiments for Lower Logistics Network Flow Distribution Model

4.1.1 Initialization

In our experiment, the PSO termination condition is reached at a pre-specified number of iterations (in our case the number of iterations was 1000). Population size is 20, the inertia weight w is decreasing linearly from 0.9 to 0.4, and acceleration (learning) factor C_1 and C_2 are set to 2.

The number of demand customer is 100. The number of intermediate node is 4, that is, there are four areas in lower logistics network. Assume that the number of demand customers in each area is 25. So transportation cost factor $c_{pj}^{md_p}$ forming intermediate node m_p to demand code d_pj is a matrix with size $4 \times 25.c_{pj}^{md_p}$ is initialized in interval [5, 10]. Capacity $y_{pj}^{md_p}$ of transport routes $L_{pj}^{md_p}$ is also a matrix with size 4×25 and is initialized in interval [5, 8]. Circulation capacity y_p^m of intermediate node m_p is a vector with size 1×4 and is initialized [200, 250]. Demand $y_j^{d_p}$ of demand nodes is a matrix with size of 4×25. Unit transportation cost is initialized to 4 Yuan, the standard deadweight of a vehicle is initialized to 2 tons.

Through the above initial value and constraints of Eq. (10), we can conclude: the decision variables $x_{pj}^{md_p}$ from intermediate node m_p to demand node d_pj can be initialized a matrix with size 4×25, and is less than $y_{pj}^{md_p}$; The decision variables x_p^m by intermediate node can be initialized by a vector with size 1×4, and is less than y_p^m.

4.1.2 Results

Lower logistics network distribution model is the multi-objective function with constraints, and converted into a single objective function with constraints based on the target weighting method. When the fitness function is respectively:

$$\text{Fitness\_low1}(x) = 0.1{*}C_{low} - 0.9{*}U, \text{Fitness\_low2}(x) = 0.2{*}C_{low} - 0.8{*}U,$$
$$\text{Fitness\_low3}(x) = 0.3{*}C_{low} - 0.7{*}U,$$
$$\text{Fitness\_low4}(x) = 0.4{*}C_{low} - 0.6{*}U, \text{Fitness\_low5}(x) = 0.5{*}C_{low} - 0.5{*}U,$$
$$\text{Fitness\_low6}(x) = 0.6{*}C_{low} - 0.4{*}U,$$
$$\text{Fitness\_low7}(x) = 0.7{*}C_{low} - 0.3{*}U, \text{Fitness\_low8}(x) = 0.8{*}C_{low} - 0.2{*}U,$$
$$\text{Fitness\_low9}(x) = 0.9{*}C_{low} - 0.1{*}U,$$

We can obtain the value of fitness function, Logistics costs and logistics network usage shown in Table 1.

The results in Table 1 show that logistics costs function and logistics network usage function assign different weight factor resulting in different objective function value. Our goal is to minimize logistics costs, to maximize logistics network usage. So the following single objective function obtains better result: Fitness_low1, Fitness_low4, Fitness_low5.

Figure 3 illustrates the convergence of objective function. The results in Fig. 3 show that with the increase of the iterations number, the objective function value is smaller, gradually moving towards convergence.

Table 1 The value of objective function, logistics cost function, logistics network usage function

	Fitness_low1	Fitness_low2	Fitness_low3	**Fitness_low4**	**Fitness_low5**
Fitness	**1.7835e + 003**	3.7587e + 003	5.6762e + 003	**7.3488e + 003**	**8.6016e + 003**
C_{low}	**18352**	19000	19088	**18568**	**17248**
U	**58.9276**	58.1167	56.6214	**58.3639**	**47.5932**

	Fitness_low6	Fitness_low7	Fitness_low8	Fitness_low9
Fitness	1.1184e + 004	1.3341e + 004	1.4492e + 004	1.5648e + 004
C_{low}	18672	19112	18480	17392
U	56.2586	58.8502	55.3139	52.2484

The bold font denote the objective function and the better results.

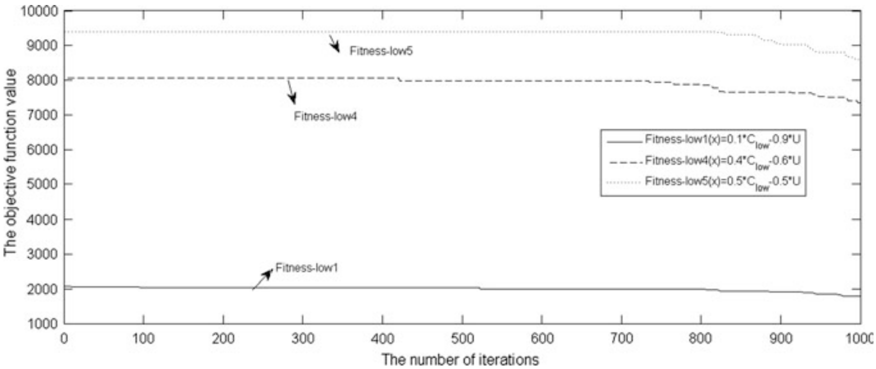


Fig. 3 The convergence of objective function

When the objective functions are respectively Fitness_low1, Fitness_low4, Fitness_low5, we can obtain the value Logistics capacity $x_{pj}^{md_p}$ from intermediate node m_p to demand node d_{pj}, which is shown in Fig. 4. $x_{pj}^{md_p}$ is less than $y_{pj}^{md_p}$ in transport routes $L_{pj}^{md_p}$.

The values of decision variables x_p^m via intermediate node are shown in Table 2. Table 2 shows that x_p^m meets the constraints condition in Eq. (10).

The running time of PSO algorithm is shown in Table 3. Table 3 shows that when the population size is 29, the number of iteration of PSO is 1000, the average running time of PSO is 64.5781 s. Therefore, based on multi-objective particle swarm optimization algorithm, we can find the solution of lower logistics network distribution model in a short period of time.

7	5	2	5	1	2	7	2	4	2	1	4	6	3	2	5	3	3	5	1	1	1	6	2	2
1	2	8	3	1	1	6	2	2	3	2	5	1	2	4	3	4	4	2	2	1	4	4	5	2
1	2	2	1	7	3	3	2	5	5	2	1	3	6	2	1	3	4	2	6	1	7	5	3	1
1	1	5	6	2	8	2	2	1	2	7	3	6	4	2	1	3	6	2	2	2	3	5	2	1

7	3	6	3	2	2	2	2	1	2	1	2	2	4	6	2	1	2	6	6	1	2	5	3	2
2	2	2	2	4	1	2	2	2	2	3	3	1	4	2	5	4	1	4	1	2	2	7	5	2
4	4	4	1	6	5	8	4	5	5	3	1	1	4	3	3	5	2	3	2	6	2	5	3	2
7	2	2	1	5	4	1	2	7	6	4	2	3	2	2	1	1	6	7	1	2	2	1	8	4

1	2	7	2	2	2	3	1	2	1	1	1	4	2	3	6	3	3	4	1	2	2	1	1	6
6	3	8	2	3	1	4	2	1	2	6	3	1	1	2	2	2	1	1	4	3	2	2	5	1
1	1	2	4	4	5	1	5	5	1	1	1	2	6	1	3	2	6	8	1	2	2	1	1	2
5	4	1	3	5	2	3	1	1	3	2	2	4	3	5	6	6	3	4	7	4	2	2	1	6

Fig. 4 The value of $x_{pj}^{md_p}$

Table 2 The value of x_p^m

Fitness	Fitness_low1	Fitness_low4	Fitness_low5
x_p^m	82 74 78 79	75 67 91 83	63 68 68 85

Table 3 The running time of PSO (unit: second)

	Fitness_low1	Fitness_low2	Fitness_low3	**Fitness_low4**	**Fitness_low5**
Run time	58.1786	66.8475	59.6456	67.5020	64.5945
	Fitness_low6	Fitness_low7	Fitness_low8	Fitness_low9	Average
Run time	70.6910	72.5468	59.2291	61.9682	64.5781

The bold font denote the objective function and the better results.

4.2 Experiment for Upper Logistics Network Flow Distribution Model

4.2.1 Initialization

The number of supply node is 20. The number of intermediate node is 4. The transportation cost factor c_{ip}^{sm} forming supply node s_i to intermediate node m_p is a matrix with size of 20 × 4 and is initialized to integer in interval [5, 10]; Inventory cost factor c_p^h via intermediate node is a vector with size of 1 × 4 and is initialized to integer in interval [5, 10]; The capacity y_{ip}^{sm} of transport routes L_{ip}^{sm} is a matrix with size of 20 × 4 and initialized to integer in interval [3, 5]. Circulation capacity y_p^h via intermediate node m_p is a vector with size of 1 × 4 and initialized to integer in interval [60, 100]. Unit transportation cost is initialized to 4 Yuan, and the standard deadweight of a vehicle is initialized to 2 tons.

Through the above initial value and constraints of Eq. (14), we can conclude: the decision variables x_{ip}^{sm} from supply node s_i to intermediate node m_p can be initialized by a matrix with size 20 × 4, and is less than y_{ip}^{sm}; The decision variables x_p^h via intermediate node to demand node can be initialized by a vector with size 1 × 4.

4.2.2 Results

Upper logistics network distribution model is also multi-objective function with constraints, and is converted into a single objective function with constraints based on the target weighting method. When the fitness function is following:

$$
\begin{aligned}
&\text{Fitness\_p1}(x) = 0.1{*}C_{\text{low}} - 0.9{*}U, \text{Fitness\_up2}(x) = 0.2{*}C_{\text{low}} - 0.8{*}U,\\
&\text{Fitness\_up3}(x) = 0.3{*}C_{\text{low}} - 0.7{*}U,\\
&\text{Fitness\_up4}(x) = 0.4{*}C_{\text{low}} - 0.6{*}U,\ \ \text{Fitness\_up5}(x) = 0.5{*}C_{\text{low}} - 0.5{*}U,\\
&\text{Fitness\_up6}(x) = 0.6{*}C_{\text{low}} - 0.4{*}U,\\
&\text{Fitness\_up7}(x) = 0.7{*}C_{\text{low}} - 0.3{*}U, \text{Fitness\_up8}(x) = 0.8{*}C_{\text{low}} - 0.2{*}U,\\
&\text{Fitness\_up9}(x) = 0.9{*}C_{\text{low}} - 0.1{*}U,
\end{aligned}
$$

We can obtain the value of fitness function, Logistics costs and logistics network usage shown in Table 4.

The results in Table 4 show that logistics costs function and logistics network usage function are assigned by different weight factor resulting in different objective function value. Our goal is to minimize logistics costs, to maximization logistics network usage. So the following single objective function obtains better results: Fitness_up3, Fitness_up4, Fitness_up8.

Table 4 The value of objective function, logistics cost function, logistics network usage function

	Fitneess_up1	Fitness_up2	**Fitness_up3**	**Fitness_up4**	Fitness_up5
Fitness	912.7811	1.9157e + 003	**2.9489e + 003**	**3.9240e + 003**	4.8138e + 003
C_{low}	9928	9928	**10040**	**9944**	9712
U	88.9099	87.3953	**90.1067**	**89.2643**	84.3981

	Fitness_up6	Fitness_up7	**Fitness_up8**	Fitness_up9
Fitness	5.9362e + 003	6.8898e + 003	**7.9120e + 003**	9.0199e + 003
C_{low}	9952	9880	**9912**	10032
U	87.5636	87.3730	**88.0700**	88.6173

The bold font denote the objective function and the better results.

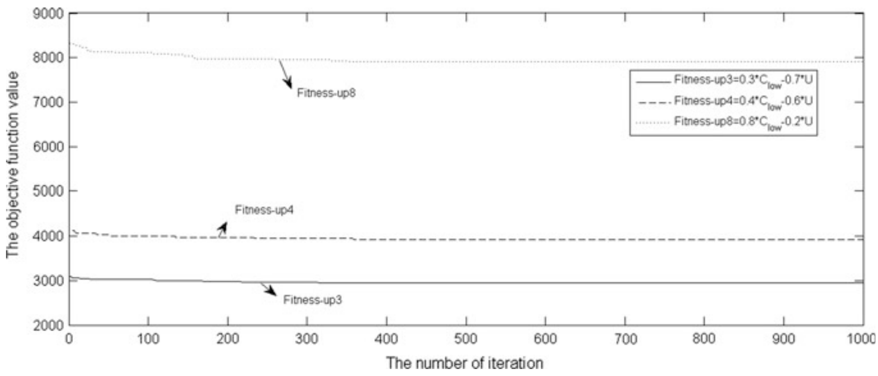


Fig. 5 The convergence of objective function

Figure 5 illustrates the convergence of objective function. The results in Fig. 5 show that with the increase number of iterations, the objective function value is smaller, gradually moving towards convergence.

The values of decision variables x_p^h are shown in Table 5. x_p^h meets the constraints condition.

The running time of PSO algorithm is shown in Table 6. Table 6 shows that when the number of people is 20, the number of iteration of PSO is 1000, and the

Table 5 The value of x_p^h

Fitness	Fitness_up3	Fitness_ up4	Fitness_up8
x_p^h	0 10 2 10	0 14 2 10	6 17 7 10

Table 6 The run time of PSO (unit: second)

	Fitness_up1	Fitness_up2	Fitness_up3	Fitness_up4	Fitness_up5
Run time	36.0037	35.3577	34.9760	35.0583	39.4409

	Fitness_up6	Fitness_up7	Fitness_up8	Fitness_up9	Average
Run time	35.0304	35.0881	53.9804	38.9201	38.2062

average running time of PSO is 38.2062 s. Therefore, based on multi-objective particle swarm optimization algorithm, we can find the solution of upper logistics network distribution model in a short period of time.

5 Conclusion

For flow distribution problem, this article design the "sliced, two-stage" flow distribution model for practical application, which consists of the lower logistics network flow distribution model and the upper logistics network flow distribution model. The model is multi-objective optimization problem with multiple constraints. We solve the two-stage model based on PSO and penalty-parameter-less method. The results of the above tables and graphs show that PSO can find logistics scale from supply node to intermediate node and from intermediate node to demand node, logistics scale and inventory of intermediate node. Viewing from the algorithm running time, we see that the model can solve the problem in about a minute, so PSO algorithm can be used to resolve flow distribution problems in a short period of time.

Acknowledgements This work was supported by the National Natural Science Fund (No. 71361008, No. 71461008), and Hainan Province key projects (No. ZDXM20130080), and the Hainan Province Natural Science Fund (No. 20151003).

References

1. Candas, M.F., Kutanoglu, E.: Benefits of considering inventory in service parts logistics network design problems with time-based service constraints. HE Trans. **39**(2), 159–176 (2007)
2. Thanh, P.N., Peton, O., Bostel, N.: A linear relaxation-based heuristic approach for logistics network design. Comput. Ind. Eng. **59**(4), 964–975 (2010)
3. Ishfaq, R., Sox, C.R.: Hub location-allocation in intermodal logistic networks. Eur. J. Oper. Res. **210**(2), 213–230 (2011)
4. Wang, B.H., He, S.W., Song, R., et al.: Multi-modal express shipment network routing optimization model and algorithm. Railway Soc. **31**(2), 12–16 (2009)
5. Ge, S., Gan, M.: Genetic algorithm based design of the supply chain network integrated with the reverse logistics. China Railway Sci. **29**(6), 116–120 (2008)
6. Dias, J., Euqenia, C.M., Climaco, J.: Capacitated dynamic location problems with opening, closure and reopening of facilities. IMA J. Manag. Math. **17**(4), 317–348 (2006)
7. Shi, Q., Chen, C.Y., Qin, Y.M.: A multi-objective optimal model on logistics network programming. Chin. J. Manag. Sci. **13**(4), 40–43 (2005)
8. Tang, X.F., Mao, H.J., Li, X.H.: Multi-objective optimization model of locations for distribution centers. J. Southeast Univ. **39**(2), 404–407 (2009)
9. Lu, X.W., Li, J., Lei, M., et al.: Multi-objective optimization of stochastic demand inventory routing problem with time windows. J. Southwest JiaoTong Univ. **44**(2), 289–294 (2009)
10. Zhan, S.L., Liu, N.: A multi-objective stochastic programming model for emergency logistics based on fuzzy goal programming. Chin. Mech. Eng. **22**(23), 2858–2862 (2011)

11. Zheng, B., Ma, Z.J., Fang, T.: Fuzzy Multi-objective location-routing problem in emergency logistics. Syst. Eng. **27**(8), 21–25 (2009)
12. Eberhart, R.C., Shi, Y.: Comparing inertia weights and constriction factors in particle swarm optimization. In: IEEE International Conference on Evolutionary Computation, pp. 81–86 (2001)
13. Deb, K.: An efficient constraint handling method for genetic algorithms. Comput. Methods Appl. Mech. Eng. **186**(2–4), 311–338 (2000)

A New Method to Obtain Deep Ocean Current Velocity from Argo Floats

Jian-guo Qian and Yong-liang Han

Abstract The ADCP (Acoustic Doppler Current Profiler) has been widely used to achieve three dimensions' current velocities directly at any layer in the ocean. However, the coverage of ADCP is limited. The widely delivery of Argo floats provides a new way to obtain the deep ocean current. To make full use of Argo floats to achieve Deep Ocean current, a new processed unit is posed here. Adaptive Kalman filter is introduced to process the time series. A numerical simulation has been carried out to validate the effect of three different analysis methods for various data source. The simulation results suggest the feasibility and stability of adaptive Kalman filter for the unit process. This method has been applied to the different depth's Argo floats, and the corresponding error estimation is given out. The accuracy has been improved dramatically compared with the unfiltered.

Keywords ADCP · Argo floats · Model · Adaptive Kalman filter · Numerical simulation

1 Introduction

To get the ocean current velocity, either the ocean surface or the deep current, we get it directly or indirectly through measurement. The three dimensions' current velocities can be achieved by ADCP (Acoustic Doppler Current Profiler) at any layer directly, and the TAO (Tropical Atmosphere Ocean) array provides us with real-time data from moored ocean buoys but few measurements are available. With the development of satellite technology, satellite altimetry has become one of the

J. Qian (✉)
College of Geomatics, Liaoning Technical University, Fuxin, Liaoning 123000, China
e-mail: 347566587@qq.com; jg_qian7529@163.com

Y. Han
College of Mining, Liaoning Technical University, Fuxin, Liaoning 123000, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_48

most useful methods to observe sea-surface velocity field globally [1]; Ki-mook Kang and Duk-jin Kim introduced a new method to extract more accurate ocean surface current velocity [2]. The satellite-tracked drifter has emerged, giving us indirectly ways to achieve the ocean current velocity [3], but it only provides us with surface velocity. Two autonomous profiling were deployed successfully in the East/Japan Sea in 1996 [4], and the Argo program has been formed in 2000. With the increasing deployment of Argo in the ocean, Argo plays a crucial role in oceanography [5], giving us a new way to study the surface ocean or the deep ocean current velocity.

To get the ocean current velocity, especially in indirect ways, the motion model of Argo or the motion model of sea water (because sea water is the carrier of Argo) must be taken into account. Jone Jin Park and Kuh Kim have developed an analysis method based on least-square principle to determine the ocean surface and deep ocean current velocity [4]. Xie and Zhu introduced Kalman filter into the estimation of the current of surface and mid-depth [6]. These two methods are under the presumption that a surface trajectory of Argo float is dominated by a constant current and inertial oscillation. However, detailed control schema must be adopted to obtain the position and time information, leading to a lot of floats out of usage. Because the position and time are just the byproduct compared with temperature and salinity information and usually are not full enough to perform detailed control. We need to provide a new method to make full use of the mass data. This paper presents a new processing method. The paper provides us with a new processing unit, extends Kalman filter method for unstable motion, and then introduces the adaptive Kalman filter into the processing of deep ocean current velocity from Argo floats. Further studies on the regional statistical analysis of deep ocean velocity will be summarized in our next study.

2 Analysis Method

Detailed analysis of Argo floats was given out by Park to obtain the deep ocean current velocity. To get the deep current velocity, the surfacing and diving positions must be extrapolated from the fix points of the sea-surface [4]. However, the position information is just the byproduct of Argo floats [7, 8]; the information is generally not enough to perform the detailed analysis. The float is under water most of the time of a cycle (Fig. 1): the average value of staying at the surface is 10.1 h and the average value of staying underwater is 9.1 days [9]. The average value of the positions during one cycle was processed as a unit in the time series to analyze the Argo floats' trajectory in deep ocean.

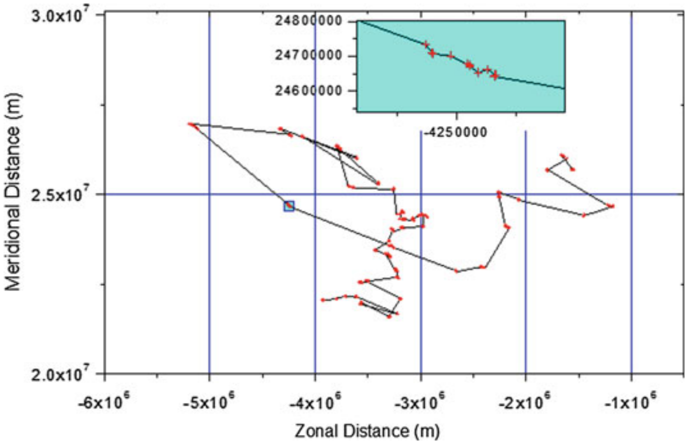


Fig. 1 Unit in the time series

2.1 Unit

Because of the lack of information to perform the detailed analysis and the underwater time's high proportion in one cycle, the unit was proposed to delegate the position in the cycle. The unit is defined as follows:

$$\begin{cases} P_i = \left(\sum_{k=1}^{n_i} P_{i_k}\right) / n_i \\ T_i = \left(\sum_{k=1}^{n_i} T_{i_k}\right) / n_i \end{cases} \tag{1}$$

where i is the index of the cycle; $P_{i_k} = (X_{i_k}, Y_{i_k})$, is the fixed point at T_{i_k} at the surface; n_i is the number of fixed points in this cycle.

Supposing that the variance of each fixed points is Q_{i_k} and the covariance between the fixed points is zero. The variance of the unit can be calculated as follows:

$$Q_i = \left(\sum_{k=1}^{n_i} Q_{i_k}\right) / n_i^2 \tag{2}$$

2.2 Motion Models

The motion of Argo floats is complicated, and three basic motion models are listed in this section.

2.2.1 CVM (Constant Velocity Model)

The formula is as follow:

$$P = P_0 + VT \tag{3}$$

The discrete form is:

$$P_k = P_{k-2} + V_{k-1}\Delta T \tag{4}$$

where$P_k = (X_k, Y_k)$, is the Argo's position at T_k; $\Delta T = T_k - T_{k-2}$; $V_{k-1} = (P_{k-1} - P_{k-2})/(T_{k-1} - T_{k-2})$.

In this model, a position can be calculated from the previous two points, but it is under the presumption that the velocity in ΔT_{i-1} is the same with ΔT_{i-1}.

2.2.2 CM (Coriolis Model)

This model is derived from Coriolis force because of earth rotation; the differential equation form is as follow:

$$\begin{cases} \frac{du}{dt} - fv = 0 \\ \frac{dv}{dt} + fu = 0 \end{cases} \tag{5}$$

where

u is the parallel circle portion of current velocity;

v is the meridian portion of current velocity;

f is the Coriolis parameter, which can be calculated from the latitude. The practical form is [4]:

$$\begin{cases} X_k = X_i + (X_0 - X_i)\cos(-f\Delta T_k) - (Y_0 - Y_i)\sin(-f\Delta T_k) \\ Y_k = Y_i + (X_0 - X_i)\sin(-f\Delta T_k) + (Y_0 - Y_i)\cos(-f\Delta T_k) \end{cases} \tag{6}$$

where

(X_i, Y_i) is the center of the inertial motion;

(X_0, Y_0) is the origin position of the inertial motion;

The recursive form is as follow:

$$\begin{cases} X_k = X_{k-2} + u_{k-1}f^{-1}\sin(f\Delta T) + v_{k-1}f^{-1}[1 - \cos(f\Delta T)] \\ Y_k = Y_{k-2} + v_{k-1}f^{-1}\sin(f\Delta T) - u_{k-1}f^{-1}[1 - \cos(f\Delta T)] \end{cases} \tag{7}$$

In this model, a position can be calculated from the previous two points, but it is under the presumption that the three points are in the same inertial motion circle.

2.2.3 CBM (Combined Model)

This model is the combination of CM and CVM, the practical form is as follow:

$$\begin{cases} X_k = X_i + (X_0 - X_i)\cos(-f\Delta T_k) - (Y_0 - Y_i)\sin(-f\Delta T_k) + u_1\Delta T_k \\ Y_k = Y_i + (X_0 - X_i)\sin(-f\Delta T_k) + (Y_0 - Y_i)\cos(-f\Delta T_k) + v_1\Delta T_k \end{cases} \tag{8}$$

In this model, at least 3 fixed points are needed to calculate the parameters; Least-square method will be used to achieve the optimal estimation for more fixed points.

The recursive form is as follow:

$$\begin{cases} X_k = X_{k-2} + \alpha u_{k-1}\Delta T_k + \beta\{u_{k-1}f^{-1}\sin(f\Delta T) + v_{k-1}f^{-1}[1 - \cos(f\Delta T)]\} \\ Y_k = Y_{k-2} + \alpha v_{k-1}\Delta T_k + \beta\{v_{k-1}f^{-1}\sin(f\Delta T) - u_{k-1}f^{-1}[1 - \cos(f\Delta T)]\} \end{cases} \tag{9}$$

In this model, a position can be calculated from the previous two points, but it is difficult to determine the weight between CM and CVM, that is, how to determine the parameters α and β. In essence, the motion of floats is conceived as the combination of linear motion and nonlinear motion. This model was adopted in the following analysis methods.

2.3 *Analysis Methods*

Considering the physical process of Argo floats in the deep ocean, CBM is adopted here. Based on the model, three different analysis methods are described as follows:

2.3.1 Least Square

As shown in formula (8), (6) unknown parameters $(u_l, v_l, X_i, Y_i, X_0, Y_0)$ are needed to be calculated to determine the motion of Argo float. The error equation is as follow:

$$\underset{n\times 1}{V} = \underset{n\times 1}{A}\ \underset{n\times 1}{\hat{x}} - \underset{n\times 1}{l} \tag{10}$$

where

$$V=\begin{pmatrix} V_{1X} \\ V_{1Y} \\ V_{2X} \\ V_{2Y} \\ \vdots \\ V_{n/2X} \\ V_{n/2Y} \end{pmatrix}, \widehat{x}=\begin{pmatrix} u_L \\ y_L \\ x_0 \\ y_0 \\ \vdots \\ x_i \\ y_i \end{pmatrix}, l=\begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_{n/2} \end{pmatrix} - A\widehat{x_0},$$

$$A=\begin{Bmatrix} \Delta t_1 & 0 & \cos(f\Delta t_1) & \sin(f\Delta t_1) & 1-\cos(f\Delta t_1) & -\sin(f\Delta t_1) \\ 0 & \Delta t_1 & -\sin(f\Delta t_1) & \cos(f\Delta t_1) & \sin(f\Delta t_1) & 1-\cos(f\Delta t_1) \\ \Delta t_2 & 0 & \cos(f\Delta t_2) & \sin(f\Delta t_2) & 1-\cos(f\Delta t_2) & -\sin(f\Delta t_2) \\ 0 & \Delta t_2 & -\sin(f\Delta t_2) & \cos(f\Delta t_2) & \sin(f\Delta t_2) & 1-\cos(f\Delta t_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta t_{n/2} & 0 & \cos(f\Delta t_{t_{n/2}}) & \sin(f\Delta t_{t_{n/2}}) & 1-\cos(f\Delta t_{t_{n/2}}) & -\sin(f\Delta t_{t_{n/2}}) \\ 0 & \Delta t_{t_{n/2}} & -\sin(f\Delta t_{t_{n/2}}) & \cos(f\Delta t_{t_{n/2}}) & \sin(f\Delta t_{t_{n/2}}) & 1-\cos(f\Delta t_{t_{n/2}}) \end{Bmatrix}$$

and $\widehat{x_0}$ is the approximate value of $\hat{x}$.

To achieve $\hat{x}$ in optimal way, the cost function $J = V^T WV$ (W is the weight matrix of the observation) must be minimized, that is:

$$\mathrm{J} = V^T WV = \min \to \frac{\partial V^T WV}{\partial \widehat{x}} = 2V^T WA = 0 \tag{11}$$

The transposed form of (11) is:

$$A^T WV = 0 \tag{12}$$

Combining formula (10) and (12), the optimal estimation of $\widehat{x}$ can be achieved:

$$\hat{x} = (A^T WA)^{-1}(A^T W)l \tag{13}$$

2.3.2 Kalman Filter

The former analysis method is under the presumption that the Argo floats' motion mode is single or the motion parameters are invariable which is impossible. To make full use of the former observation data and the real time data, Kalman filter is adopted here. To make full use of the former information to predict the following position, the previous two points are used in the formula (14).

$$\begin{cases} \bar{X}_k = X_{k-2} + \alpha u_{k-1} \Delta T_k + \beta \{ u_{k-1} f^{-1} \sin(f \Delta T) + v_{k-1} f^{-1} [1 - \cos(f \Delta T)] \} \\ \bar{Y}_k = Y_{k-2} + \alpha v_{k-1} \Delta T_K + \beta \{ v_{k-1} f^{-1} \sin(f \Delta T) - u_{k-1} f^{-1} [1 - \cos(f \Delta T)] \} \end{cases} \tag{14}$$

where

$(\overline{X_k}, \overline{Y_k})$ is the predicted position;

(u_{k-1}, v_{k-1}) can be calculated from the previous two points, that is:

$$u_{k-1} = (X_{k-1} - X_{k-2})/\Delta T_{k-2,k-1} \text{ and } v_{k-1} = (Y_{k-1} - Y_{k-2})/\Delta T_{k-2,k-1}.$$

In fact, the formula is the discrete form of the system equation.

In other hand, the position can be achieved from the direct observation. The observation equation is:

$$\begin{cases} X_K = X_{\text{ok}} \\ Y_k = Y_{ok} \end{cases} \tag{15}$$

The error equation form of system and observation equations is:

$$\begin{cases} V_{\overline{P_k}} = \widehat{p_k} - \overline{p_k} \\ V_K = \widehat{p_k} - p_{ok} \end{cases} \tag{16}$$

where

$\overline{p_k} = (\overline{X_k}, \overline{Y_k})$, which can be achieved from formula (14) and the covariance can be obtained; $P_{0k} = (X_{0k}, Y_{0k})$, is the observation.

To achieve $\widehat{p_k}$ in optimal way, the cost function $J(k) = V_k^T W_k V_k + V_{\overline{P_k}}^T W_{\overline{P_k}} V_{\overline{P_k}}$. ($W_k$ is the weight matrix of the observation and $W_{\overline{P}_k}$ is the weight matrix of the predicted value) must be minimized, that is:

$$\frac{dJ(k)}{dp_k} = 2V_k^T W_k + 2V_{\overline{P}_k}^T W_{\overline{P}_k} = 0 \tag{17}$$

Combining formula (17) and (12), the optimal estimation of $\widehat{p_k}$ can be achieved:

$$\widehat{p_k} = \left(W_k + W_{\overline{P_k}} \right)^{-1} \left(W_k P_{ok} + W_{\overline{P_k}} \overline{P_k} \right) \tag{18}$$

And the covariance of $\widehat{p_k}$ also can be obtained for the next filtering process.

2.3.3 Adaptive Kalman Filter

Kalman filter takes the former observation and the real time observation into account, but the weight between former and real time information is equal.

Adaptive Kalman filter is adopted here to adjust the weight between former and real time information in a reasonable way. The principle of adaptive Kalman filter is similar to Kalman filter. The cost function of adaptive Kalman filter is:

$$J(k) = V_k^T W_k V_k + \eta_k V_{\overline{P_k}}^T W_{\overline{P_k}} V_{\overline{P_k}} \tag{19}$$

where

η is the adaptive factor, which can be calculated from the previous information;

$$\frac{dJ(k)}{d\widehat{p_k}} = 2V_k^T W_k + 2\eta_k V_{\overline{P_k}}^T W_{\overline{p_k}} = 0 \tag{20}$$

Combining formula (20) and (12), the optimal estimation of $\widehat{p_k}$ can be achieved:

$$\widehat{p_k} = (W_k + \eta_k W_{\overline{P_k}})^{-1}(W_k P_{0k} + \eta_k W_{\overline{P_k}} \overline{P_k}) \tag{21}$$

3 Result

3.1 Numerical Simulation

Because of the variability of the Argo float in Deep Ocean, a numerical simulation was performed to evaluate the effect of the three analysis methods. All the data here was generated based on the presumption that the motion of Argo floats was the combination of linear motion and inertial motion. So the motion or the data was decided by the linear parameters (u_l, v_l) and the inertial parameters (X_i, Y_i, X_0, Y_0). Four kinds of data were generated here: constant linear parameters and constant inertial parameters (CLCI), variable linear parameters and constant inertial parameters (VLCI), constant linear parameters and variable inertial parameters (CLVI), variable linear parameters and variable inertial parameters (VLVI). As shown in Fig. 2.

The sixth point is the turning point, after this the linear or the inertial parameters are changed and the corresponding trajectories are changed. These four kinds of data are the basic simulation data for the following analysis.

First of all, the least-square method is applied to obtain the motion parameters, including the linear parameters and inertial parameters, and then the parameters are used to get the position series. The effect of least-square method is shown in Fig. 2. At the turning point, CLCI presents stable attribute, and the others present unstable attribute not only at turning point but also in the total time series. The departure of the calculated results from the actual data is listed in Fig. 2. The departure of VLVI is obvious greater than others, and the CLCI departure is just around 1000 m in a stable way. To achieve better results for all kinds of data, the Kalman filter method is applied. The effect of Kalman filter is shown in Fig. 3. Compared with least-square method, the Kalman filter is more accuracy. The departure of all kinds

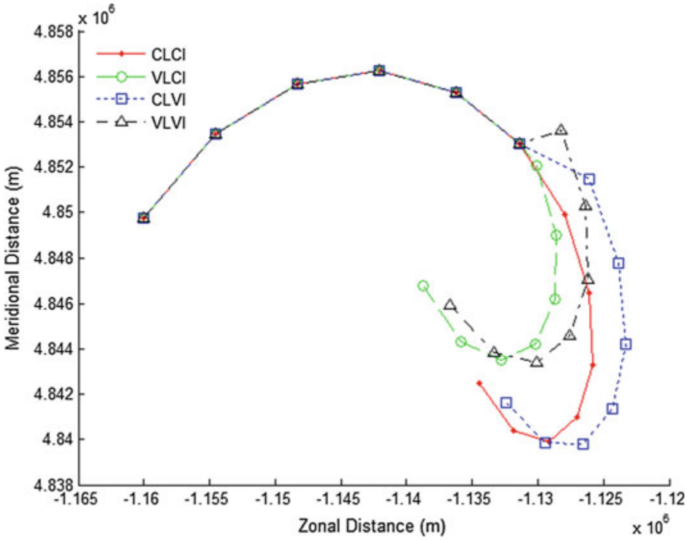


Fig. 2 Four kinds of simulation data

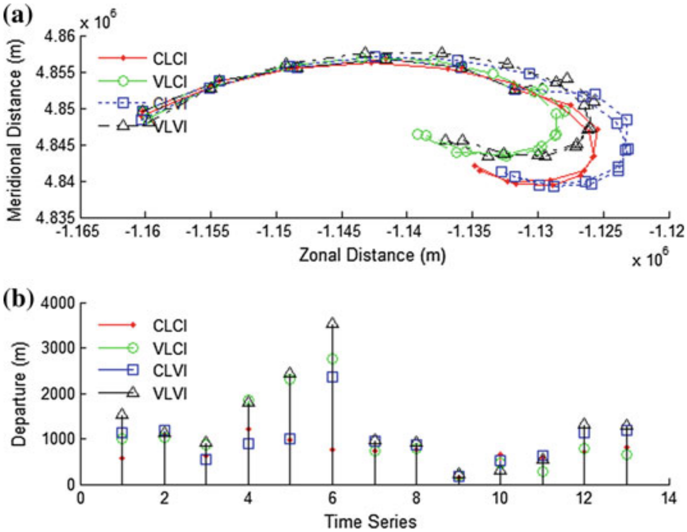


Fig. 3 The effects of least-square method in various data

of data is smaller than the correspondence of the previous method. The departure of VLVI is also greater than others. The effect of adaptive Kalman filter is slightly better than Kalman filter. Table 1 shows the standard deviation (STD) of different time series for different methods. The STD of CLCI is the smallest and the STD of VLVI is the greatest. The STD of VLVI is greater than CLVI and VLCI, which

Table 1 Standard deviation (m)

	Data method			
	CLCI	VLCI	CLVI	VLVI
Least-square	817	1343	1131	1629
Kalman filter	362	424	463	515
Adaptive Kalman filter	358	420	459	510
Adaptive Kalman filter	358	420	459	510

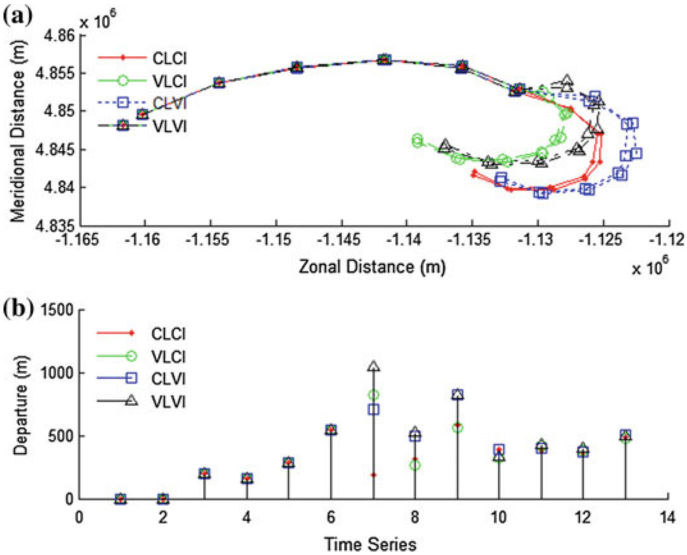


Fig. 4 The effects of Kalman filter in various data

suggests that the complicated level of motion is a crucial factor for the least-square method. Adaptive Kalman filter can achieve better results for all kinds of data (Figs. 4 and 5).

3.2 *The Deep Ocean Current Velocity from Argo Floats*

Adaptive Kalman Filter is adopted here to calculate the unit position, three floats of different parking depth are calculated to determine the trajectory of floats, and then the current velocity from the trajectory. The ID of floats are 69036, 5900764, and 5900768. The parking depth are 400 m, 1000 m, and 1750 m respectively. The accuracy of the estimation unit P can be calculated from the formula (21). The

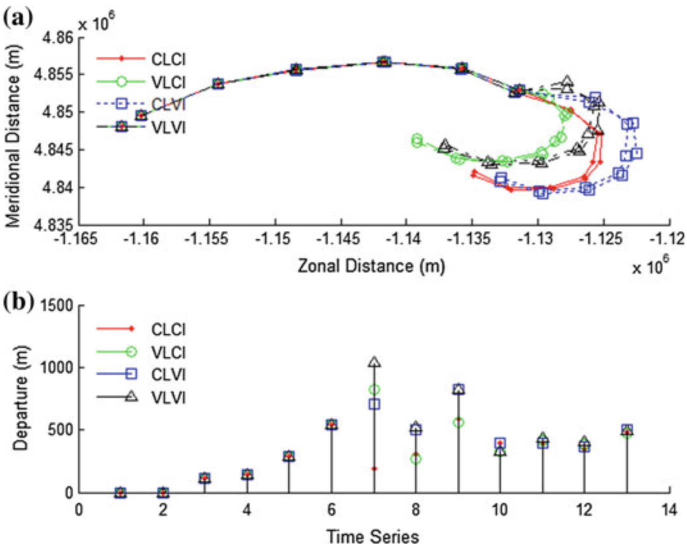


Fig. 5 The effects of adaptive Kalman filter in various data

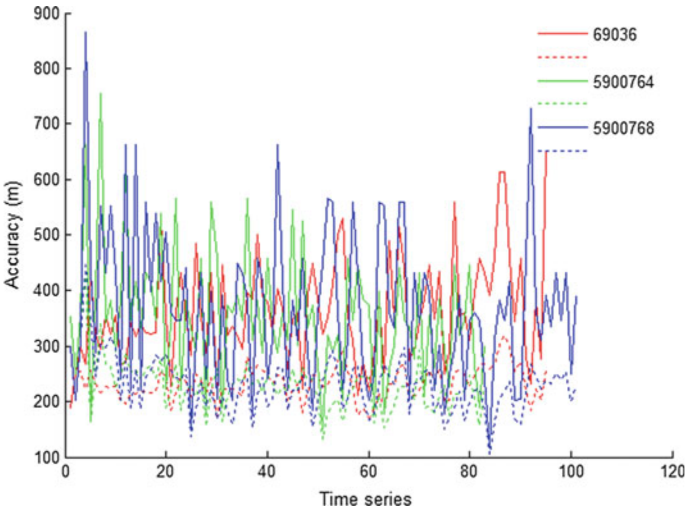


Fig. 6 The accuracy of filtered and unfiltered unit

accuracy of filtered and unfiltered is shown in Fig. 6, The solid line is the unfiltered accuracy and the dash line is the filtered accuracy. From the three groups of data; we can come into a conclusion that through the adaptive Kalman filter method the accuracy of the position of the unit is dramatically improved.

Position of the unit is determined, and the current velocity can be calculated as follow:

$$V_k = (P_{k+1} - P_k)/\Delta T_k \tag{22}$$

And the accuracy of V_k can be derived from the accuracy of P_{k+1} and P_k.

4 Conclusion

To make full use of Argo floats to determine the deep ocean current velocity, the processing unit is defined here. Three different motion models are presented; the combined model is adopted for the analysis. In comparing the unit position obtained from the least-square method, Kalman filter, and adaptive Kalman filter we analyze their effects to various data source. The results show that the Least-square method is sensitive to the variability of motion model; the Kalman filter can achieve stable result for various data source; and the adaptive Kalman filter can achieve perfect result. The adaptive Kalman filter is applied to Argo floats with different parking depth to obtain the unit position, and then the current velocity. The corresponding error estimations are presented. We come into a conclusion that the accuracy of unit position is dramatically improved through adaptive Kalman filter. The application of this method for the regional research of current will be presented in the following research.

References

1. Uchida, H., Imawaki, S., Hu, J.H.: Comparison of Kuroshio surface velocities derived from satellite altimeter and drifting buoy data. J. Oceanogr. **54**(1), 115–122 (1998)
2. Kang, K.M., Kim, D.J.: Estimation of ocean surface velocity in tropical cyclones using RADARSAT-1 ScanSAR raw data. Synthetic Aperture Radar (APSAR). In: 2011 3rd International Asia-Pacific Conference on IEEE (2011)
3. Idris, N., Mohd, M.I.S., Cracknell, A.P., et al.: Studies on the sea surface current circulation pattern in the South China Sea derived from satellite altimetry. J. Sustain. Sci. Manag. **4**(2), 27–35 (2009)
4. Park, J.J., Kim, K., Crawford, W.R.: Inertial currents estimated from surface trajectories of ARGO floats. Geophys. Res. Lett. **31**(13), 137–151 (2004)
5. Freeland, H.J.: Vertical velocity estimates in the North Pacific using Argo floats. Deep Sea Res. Part II **85**(4), 75–80 (2013)
6. Xie, J.P., Zhu, J.: Estimation of the surface and mid-depth currents from Argo floats in the Pacific and error analysis. J. Mar. Syst. **73**(1–2), 61–75 (2008)
7. Liu, H.W., Zhang, Q.L., Duan, Y.L. et al.: The three-dimensional structure and seasonal variation of the North Pacific meridional overturning circulation. Acta Oceanol. Sinica **30**(3), 33–42 (2011)

8. Stryker, S.A., DiMarco, S.F., Stoessel, M.M. et al.: Deep and surface circulation in the Northwest Indian Ocean from Argo, surface drifter, and in situ profiling current observations. American Geophysical Union, Fall Meeting Abstracts (2010)
9. Lebedev, K.V., Yoshinari, H., Maximenko, N.A., et al.: velocity data assessed from trajectories of Argo floats at parking level and at the sea surface. IPRC Tech. Note **4**(2), 1–16 (2007)

Quantum State Sharing for Networks Using the GHZ Channel

Jun Liu, Zhi-wen Mo and Shu-qin Sun

Abstract In this paper, we investigate the quantum state sharing scheme for the satellite-to-ground quantum communication network using the GHZ channel. A single-qubit state network communication scheme is presented. Through entanglement swapping, the teleportation scheme is feasible. The probability of successful transition is 1. For a two-particle qubit state, we consider a sharing state scheme in the framework of the satellite-to-ground quantum communication network using the GHZ channel. In this scheme Alice has to share a quantum state with Bob and Charlie through networks such that the collaboration of Bob and Charlie can reconstruct the quantum state, while one of them cannot obtain anything about the quantum state. Through carrying out Bell measurements, local von Neumann measurements, appropriate unitary operations, we can obtain perfect the success probability of sharing two-particle quantum state.

Keywords Quantum entanglement · Entanglement swapping · Communication network

1 Introduction

Secret sharing is one of the useful tools in the cryptographic applications. Quantum secret sharing is the generalization of classical secret sharing to the quantum scenario, and it has attracted much attention since 1999 [1, 2]. The original quantum secret sharing scheme, Hillery et al. originally considered quantum secret sharing via three-particle and four-particle entangled Greenberger-Horne-Zeilinger

J. Liu (✉) · Z. Mo
College of Mathematics and Software Science, Sichuan Normal University, Chengdu 610066, China
e-mail: liujun8476@163.com

S. Sun
Department of Mathematics, Sichuan Minzu College,
Kangding 626001, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_49

(GHZ) states. Since the publication of this pioneering work, several variations and theoretical expansions of quantum secret sharing have been reported. Recently, some research [3–5] has focused on quantum secret sharing because of its potential application in quantum information theory. Some developments have shown that both classical information and quantum information can be shared with quantum mechanics.

Quantum resources allow secret sharing to be extended in two ways: the security of classical information sharing in crypto-communication systems [6, 7]; and they can extend secret sharing to include the dissemination of quantum states [8–10] in the context of quantum information science. The second class of protocols, which we term quantum state sharing, have broad ramifications across quantum information science. Teleported states, quantum computer output states, and quantum keys used for quantum cryptography can all be securely distributed using quantum state sharing. Furthermore quantum state sharing can be used to distribute entanglement over distances and through unreliable channels, and is an enabling step towards quantum error correction.

However, litter of the existing quantum state sharing protocols have considered in the networks case. So, how to share a quantum state in the network models? It is very interesting to investigate the quantum state sharing [11] scheme for networks. In this paper, we investigate the quantum state sharing scheme for the satellite-to-ground quantum communication network [12] using the GHZ channel [13]. A single-qubit state network communication scheme is presented. Through entanglement swapping [14, 15], the teleportation scheme is feasible. The probability of successful transmission is 1. For a two-particle qubit state, we consider a sharing state scheme in the framework of the satellite-to-ground quantum communication network using the GHZ channel. In this scheme Alice has to share a quantum state with Bob and Charlie through networks such that the collaboration of Bob and Charlie can reconstruct the quantum state, while one of them cannot obtain anything about the quantum state. Through carrying out Bell measurements, local von Neumann measurements, appropriate unitary operation, we can obtain perfect the success probability of sharing two-particle quantum state.

2 The GHZ Channel and Entanglement Swapping

Let us begin with a brief review of quantum channel using the GHZ state and entanglement swapping. In this paper, a three particle entanglement is used as quantum channel. There is a complete set of the three particle entangled states of the form

$$\begin{aligned} &1/\sqrt{2}[|000\rangle \pm |111\rangle], 1/\sqrt{2}[|001\rangle \pm |110\rangle], \\ &1/\sqrt{2}[|010\rangle \pm |101\rangle], 1/\sqrt{2}[|011\rangle \pm |100\rangle]. \end{aligned} \tag{1}$$

They can be transformed into each other through unitary operations. Without loss of generality, a triplet of the GHZ form can be chosen as the quantum channel shared multi users, qubit 1, qubit 2 and qubit 3,

$$|\Phi\rangle_{123} = 1/\sqrt{2}[|000\rangle + |111\rangle]. \tag{2}$$

Consider the initial GHZ state of two pair of particles 1, 2, 3 and 4, 5, 6 to be

$$|\Phi\rangle_{123}(456) = 1/\sqrt{2}[|000\rangle + |111\rangle]. \tag{3}$$

Entanglement swapping can readily be written as

$$\begin{aligned}|\Phi\rangle_{123} \otimes |\Phi\rangle_{456} = 1/2|\Phi^{\pm}\rangle_{36} \otimes 1/\sqrt{2}(|0000\rangle \pm |1111\rangle)_{1245}\\ + 1/2|\Psi^{\pm}\rangle_{36} \otimes 1/\sqrt{2}(|0011\rangle \pm |1100\rangle)_{1245}\end{aligned} \tag{4}$$

where the four Bell states are defined by

$$|\Phi^{\pm}\rangle = 1/\sqrt{2}(|00\rangle \pm |11\rangle), |\Psi^{\pm}\rangle = 1/\sqrt{2}(|01\rangle \pm |10\rangle). \tag{5}$$

That is, particles 1, 2 and 3 are mutually entangled (in a GHZ state), and particles 4, 5 and 6 are mutually entangled (also in a GHZ state). When we conduct a measurement of the Bell operator on particles 3 and 6, then the joint state of the four particles become either of the following four:

$$\begin{aligned}&|\Phi^{\pm}\rangle_{36} \otimes 1/\sqrt{2}(|0000\rangle \pm |1111\rangle)_{1245}\\ &|\Psi^{\pm}\rangle_{36} \otimes 1/\sqrt{2}(|0011\rangle \pm |1100\rangle)_{1245}\end{aligned} \tag{6}$$

No matter what the outcome is, the particles 1, 2, 4 and 5 are now in one of four-particle entanglement GHZ states. Whereas prior to the measurement, the GHZ pairs were (1, 2, 3) and (4, 5, 6), after the measurement the GHZ state is (1, 2, 4, 5). It can easily be shown that the same fact would hold true even if (1, 2, 3) and (4, 5, 6) started in some other GHZ states. A pictorial way of representing the above process is given in Fig. 1.

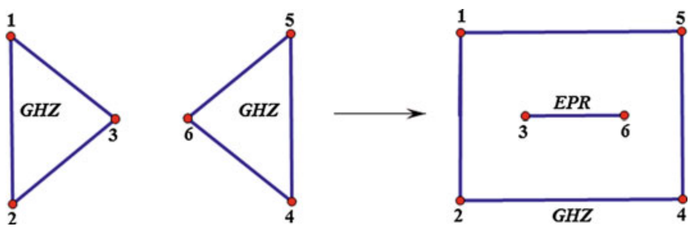


Fig. 1 The swapping of entanglement between pairs of GHZ-state particles due to a Bell state measurement on two of them

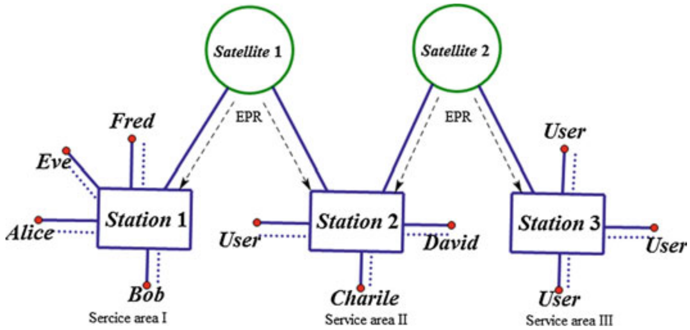


Fig. 2 Structure of satellite-to-ground quantum communication network. The *dashed* and *solid lines* represent quantum and classical wireless channels respectively

3 Network Model

In this section, we consider a satellite-to-ground quantum communication network, which is described in [12]. Some change is made in our paper. We suppose the satellite ground station shares a GHZ state ($|\Phi\rangle_{123}$) between itself and any another two quantum communication terminals that are inside its same service area. The two involved satellite ground stations share a Bell state ($|\Phi^+\rangle$). The structure of satellite-to-ground quantum communication network is shown in Fig. 2.

4 Single-Qubit State Network Communication Scheme

In this section, we discuss the single-qubit state quantum network communication scheme. In the satellite-to-ground quantum communication network, suppose that the target quantum state we seek to transmit is $|\varphi\rangle_X = \alpha|0\rangle + \beta|1\rangle$, where $|\alpha|^2 + |\beta|^2 = 1$. A quantum terminal Alice intends to transfer $|\varphi\rangle_X$ to another terminal.

If terminals Alice and receiver Bob are in the same service area, then Alice and Bob can share a GHZ state $|\Phi\rangle_{123}$ with satellite ground station, where qubit X and 1 are hold by Alice, qubit 2 by Station, and qubit 3 by Bob respectively. Using the decomposition into Bell states as before, we can rewrite the initial product state $|\varphi\rangle_X \otimes |\Phi\rangle_{123}$ as

$$\begin{aligned}|\varphi\rangle_X \otimes |\Phi\rangle_{123} = & 1/2\sqrt{2}[|\Phi^+\rangle \otimes \pi^\pm \otimes (\alpha|0\rangle \pm \beta|1\rangle) + |\Phi^-\rangle \otimes \pi^\pm \otimes (\alpha|0\rangle \mp \beta|1\rangle) \\ & + |\Psi^+\rangle \otimes \pi^\pm \otimes (\alpha|1\rangle \pm \beta|0\rangle) + |\Psi^-\rangle \otimes \pi^\pm \otimes (\alpha|1\rangle \mp \beta|0\rangle)]\end{aligned} \tag{7}$$

In order to teleport $|\varphi\rangle_X$, Alice performs a Bell basis measurement on her particles (X, 1). Station perform a von Neumann measurement ($\pi^+ = |0\rangle + |1\rangle/\sqrt{2}, \pi^- = |0\rangle - |1\rangle/\sqrt{2}$). Afterwards they publish her outcome via classical channel. At the

Table 1 Relation between measurement results and local unitary operations

Alice and Station's measurement results	Bob's local unitary operations
$\lvert\Phi^+\rangle \otimes \pi^+, \lvert\Phi^-\rangle \otimes \pi^-$	I
$\lvert\Phi^+\rangle \otimes \pi^-, \lvert\Phi^-\rangle \otimes \pi^+$	σ_Z
$\lvert\Psi^+\rangle \otimes \pi^+, \lvert\Phi^-\rangle \otimes \pi^-$	σ_X
$\lvert\Psi^+\rangle \otimes \pi^-, \lvert\Phi^-\rangle \otimes \pi^+$	$i\sigma_Y$

same time, the state of particle 3 with Bob is collapsed into one of the following states

$$\alpha|0\rangle \pm \beta|1\rangle, \alpha|1\rangle \pm \beta|0\rangle. \tag{8}$$

According to Alice's Bell-state measurement, and Station's von Neumann measurement result, Bob performs a proper unitary operation on particle 3, as shown in Table 1 (relation between Alice and Station's measurement results and Bob's local unitary operations). After unitary operations, Bob can recover the target qubit $|\varphi\rangle_X$. In fact, regardless of the outcome, the qubit transmission is feasible.

If terminals are not in the same service area, such as sender Alice in service area I, receiver David in service area II, then an appropriate quantum path must be determined. Alice → station 1 → station 2 → David is assumed to be the selected quantum path. According to previous assumptions, Alice and Bob share a GHZ state ($|\Phi\rangle_{AB1} = 1/\sqrt{2}[|000\rangle + |111\rangle] = 1/\sqrt{2}[|0\rangle^{\otimes 3} + |1\rangle^{\otimes 3}]$, where qubit A is hold by Alice, qubit B by Bob and qubit 1 by Station 1 respectively.) Station 1 shares a Bell state $|\Phi^+\rangle_{23}$ with Station 2 (where qubit 2 is hold by Station 1, qubit 3 by Station 2 respectively). Charlie and Daive share a GHZ state ($|\Phi\rangle_{CD4} = 1/\sqrt{2}[|0\rangle^{\otimes 3} + |1\rangle^{\otimes 3}]$ with Station 2, where qubit C is hold by Charlie, qubit D by David and qubit 4 by Station 2 respectively). At the first step, Station 1 needs to make a Bell-state measurement on particles (1, 2), and Station 2 needs to make a Bell-state measurement on particles (3, 4). Then communication channel $\Omega = |\Phi\rangle_{AB1} \otimes |\Phi^+\rangle_{23} \otimes |\Phi\rangle_{CD4}$ can be written as

$$\begin{aligned}\Omega = 1/4\sqrt{2}\{&|\Phi_{12}^+\rangle \otimes [|\Phi_{34}^\pm\rangle \otimes (|0\rangle^{\otimes 4} \pm |1\rangle^{\otimes 4}) + |\Psi_{34}^\pm\rangle \otimes (|0011\rangle \pm |1100\rangle)]\\ &+ |\Phi_{12}^-\rangle \otimes [|\Phi_{34}^\pm\rangle \otimes (|0\rangle^{\otimes 4} \mp |1\rangle^{\otimes 4}) + |\Psi_{34}^\pm\rangle \otimes (|0011\rangle \mp |1100\rangle)]\\ &+ |\Psi_{12}^+\rangle \otimes [|\Phi_{34}^\pm\rangle \otimes (|1100\rangle \pm |0011\rangle) + |\Psi_{34}^\pm\rangle \otimes (|1\rangle^{\otimes 4} \pm |0\rangle^{\otimes 4})]\\ &+ |\Psi_{12}^-\rangle \otimes [|\Phi_{34}^\pm\rangle \otimes (-|1100\rangle \mp |0011\rangle) + |\Psi_{34}^\pm\rangle \otimes (-|1\rangle^{\otimes 4} \mp |0\rangle^{\otimes 4})]\}\end{aligned} \tag{9}$$

According to Station 1 and 2's Bell-state measurement results, the state of communication channel is collapsed into one of the following states.

$$1/\sqrt{2}(|0000\rangle \pm |1111\rangle), 1/\sqrt{2}(|0011\rangle \pm |1100\rangle). \tag{10}$$

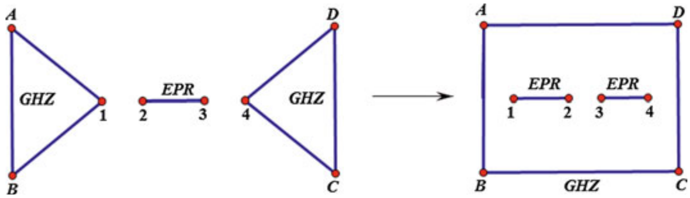


Fig. 3 The swapping of entanglement between pairs of GHZ-state particles and an EPR pair due to two Bell states measurement on them

Table 2 Relation between measurement results and communication channel

Station 1, 2's measurement	Bob's measurement	States of channel
$\lvert\Phi_{12}^{\pm}\rangle \otimes \lvert\Phi_{34}^{\pm}\rangle; \lvert\Psi_{12}^{\pm}\rangle \otimes \lvert\Psi_{34}^{\pm}\rangle$	π^+	$1/\sqrt{2}(\lvert 000\rangle + \lvert 111\rangle)$
	π^-	$1/\sqrt{2}(\lvert 000\rangle - \lvert 111\rangle)$
$\lvert\Phi_{12}^{\pm}\rangle \otimes \lvert\Phi_{34}^{\mp}\rangle; \lvert\Psi_{12}^{\pm}\rangle \otimes \lvert\Psi_{34}^{\mp}\rangle$	π^+	$1/\sqrt{2}(\lvert 000\rangle - \lvert 111\rangle)$
	π^-	$1/\sqrt{2}(\lvert 000\rangle + \lvert 111\rangle)$
$\lvert\Phi_{12}^{\pm}\rangle \otimes \lvert\Psi_{34}^{\pm}\rangle; \lvert\Psi_{12}^{\pm}\rangle \otimes \lvert\Psi_{34}^{\pm}\rangle$	π^+	$1/\sqrt{2}(\lvert 011\rangle + \lvert 100\rangle)$
	π^-	$1/\sqrt{2}(\lvert 011\rangle - \lvert 100\rangle)$
$\lvert\Phi_{12}^{\pm}\rangle \otimes \lvert\Phi_{34}^{\mp}\rangle; \lvert\Psi_{12}^{\pm}\rangle \otimes \lvert\Psi_{34}^{\mp}\rangle$	π^+	$1/\sqrt{2}(\lvert 011\rangle - \lvert 100\rangle)$
	π^-	$1/\sqrt{2}(\lvert 011\rangle + \lvert 100\rangle)$

The process schematic is shown in Fig. 3. After the Bell-state measurement, we perform a von Neumann measurement ($\pi^+ = \lvert 0\rangle + \lvert 1\rangle/\sqrt{2}, \pi^- = \lvert 0\rangle - \lvert 1\rangle/\sqrt{2}$) on particle B (or C). Then the state of communication channel transforms into a GHZ channel. The relation between Station 1, 2's Bell-state measurement results and states of communication channel is shown in Table 2.

According to previous discussion, single quantum state can be transferred by GHZ channel for the satellite-to-ground quantum communication network.

5 Two-Particle State Sharing Communication Networks Scheme

In this section, we discuss the two-particle qubit state sharing for quantum network communication scheme. In the satellite-to-ground quantum communication network, suppose that the terminal Alice (sender) wants to share the quantum state $\lvert\varphi\rangle_{XY} = \alpha\lvert 01\rangle + \beta\lvert 10\rangle$ with another two terminals which are spatial separately, where $\lvert\alpha\rvert^2 + \lvert\beta\rvert^2 = 1$.

If terminals Alice and another two share parties (Bob and Eve) are in the same service area (Service area I), then Alice and Fred share a GHZ state ($\lvert\Phi\rangle_{AF1} = 1/\sqrt{2}[\lvert 000\rangle + \lvert 111\rangle]$ with Station 1, where qubit A is hold by Alice, qubit F by Fred and qubit 1 by Station 1 respectively). Bob and Eve share a GHZ state ($|\Phi\rangle_{BE2} = 1/\sqrt{2}[|000\rangle + |111\rangle]$ with Station 1, where qubit B is hold by Bob, qubit E by Eve and qubit 2 by Station 2 respectively), according to previous assumptions. The process can refer to Fig. 1. At the first step, Station 1 needs to make a Bell-state measurement on particles (1, 2). Then the state of quantum channel becomes a four-particle entanglement GHZ state which is one of following states:

$$1/\sqrt{2}(|0000\rangle + |1111\rangle)_{AFBE}, 1/\sqrt{2}(|0011\rangle + |1100\rangle)_{AFBE}. \quad (11)$$

When the measurement result is $|\Phi^+\rangle_{12}$, the product state $\Gamma = |\varphi\rangle_{XY} \otimes 1/\sqrt{2}(|0\rangle^{\otimes 4} + |1\rangle^{\otimes 4})_{AFBE}$ can be written as

$$\begin{aligned}\Gamma =& 1/2\sqrt{2}[\pi^+ \otimes [|\Phi^\pm\rangle \otimes (\pm\alpha|1\rangle^{\otimes 3} + \beta|0\rangle^{\otimes 3}) + |\Psi^\pm\rangle \otimes (\pm\alpha|0\rangle^{\otimes 3} + \beta|1\rangle^{\otimes 3})]] \\ &+ 1/2\sqrt{2}[\pi^- \otimes [|\Phi^\pm\rangle \otimes (\pm\alpha|1\rangle^{\otimes 3} - \beta|0\rangle^{\otimes 3}) + |\Psi^\pm\rangle \otimes (\pm\alpha|0\rangle^{\otimes 3} - \beta|1\rangle^{\otimes 3})]]\end{aligned} \quad (12)$$

At the second step, Alice performs a von Neumann measurement ($\pi^+ = |0\rangle + |1\rangle/\sqrt{2}, \pi^- = |0\rangle - |1\rangle/\sqrt{2}$) on particle X and a Bell basis measurement on her particles (Y, A). Afterwards they publish their outcomes via classical channel. At the same time, the state of particles Fred, Bob and Eve are collapsed into one of the following states

$$\begin{aligned}&1/\sqrt{2}(\pm\alpha|111\rangle + \beta|000\rangle), 1/\sqrt{2}(\pm\alpha|000\rangle + \beta|111\rangle) \\ &1/\sqrt{2}(\pm\alpha|111\rangle - \beta|000\rangle), 1/\sqrt{2}(\pm\alpha|000\rangle - \beta|111\rangle)\end{aligned} \quad (13)$$

According to Alice's measurement result, Fred performs a von Neumann measurement ($\pi^+ = |0\rangle + |1\rangle/\sqrt{2}, \pi^- = |0\rangle - |1\rangle/\sqrt{2}$) on particle F. Bob and Eve perform a proper unitary operation on particles B and E, as shown in Table 3 (Relation between Alice and Fred's measurement results and local unitary operations). After unitary operations, Alice shares the quantum state $|\varphi\rangle_{XY} = \alpha|01\rangle + \beta|10\rangle$ with Bob and Eve which are spatial separately. In fact, regardless of the outcome, the sharing state scheme is feasible.

Table 3 Relation between measurement results and local unitary operations

Alice's measurement	Fred's measurement	Bob and Eve's unitary operations
$\pi^\pm \otimes |\Phi^\pm\rangle$	$\pi^\pm$	$\sigma_X \otimes I, i\sigma_Y \otimes I$
$\pi^\pm \otimes |\Phi^\mp\rangle$	$\pi^\pm$	$i\sigma_Y \otimes I, \sigma_X \otimes I$
$\pi^\pm \otimes |\Psi^\pm\rangle$	$\pi^\pm$	$I \otimes \sigma_X, I \otimes i\sigma_Y$
$\pi^\pm \otimes |\Psi^\mp\rangle$	$\pi^\pm$	$I \otimes i\sigma_Y, I \otimes \sigma_X$

If terminals Alice and sharing parties are not in the same service area, such as sender Alice in service area I, sharing parties Charlie and David in service area II, then an appropriate quantum path must be determined. Alice → station 1 → station 2 → Charlie and David is assumed to be the selected quantum path. According to previous discussion, the communication channel can be transformed to a four-particle entanglement GHZ state using entanglement swapping. So we can utilize above scheme to achieve sharing two-particle quantum state.

6 Conclusion

In the framework of the satellite-to-ground quantum communication network using the GHZ channel, we consider the teleportation scheme and sharing quantum state scheme. These schemes are feasible. Through carrying out Bell measurement, local von Neumann measurement, appropriate unitary operations, we can obtain perfect the success probability of sharing two-particle quantum state.

Acknowledgments This work is supported by Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20135134110003), Scientific Reserch Fund of SiChuan Provincial Education Department (Grant No. 11ZB153, 2015JY0002) and Scientific Research Fund of Sichuan Minzu College (Grant No. 13XYZB011, 12XYZB006).

References

1. Hillery, M., Buzek, V., Berthiaume, A.: Phys. Rev. A **59**, 1829 (1999)
2. Cleve, R., Gottesman, D., Lo, H.-K.: Phys. Rev. Lett. **83**, 648 (1999)
3. Li, Q., Chan, W.H., Long, D.Y.: Semiquantum secret sharing using entangled states. Phys. Rev. A **82**(2), 022303 (2010)
4. Zhang, Z., Li, Y., Man, Z.: Multiparty quantum secret sharing. Phys. Rev. A **71**(4), 044301 (2005)
5. Deng, F.G., Zhou, H.Y., Long, G.L.: Bidirectional quantum secret sharing and secret splitting with polarized single photons. Phys. Lett. A **337**(4), 329–334 (2005)
6. Karlsson, A., Koashi, M., Imoto, N.: Phys. Rev. A **59**, 162 (1999)
7. Tittel, W., Zbinden, H., Gisin, N.: Phys. Rev. A **63**, 042301 (2001)
8. Tyc, T., Sanders, B.C.: Phys. Rev. A **65**, 42310 (2002)
9. Tyc, T., Rowe, D.J., Sanders, B.C.: J. Phys. A: Math. Gen. **36**, 7625 (2003)
10. Lance, A.M., Symul, T., Bowen, W.P., Sanders, B.C., Lam, P.K.: quant-ph//0311015 (2003)
11. Lance, A.M., Symul, T., Bowen, W.P., et al.: Quantum state sharing. In: Second International Symposium on Fluctuations and Noise. International Society for Optics and Photonics, pp. 100–109 (2004)
12. Na, C., Dong-xiao, Q., Chang-xing, P.: Quantum communication for satellite-to-ground networks with partially entangled states. Chin. Phys. B **24**, 020304 (2015)
13. Gorbachev, V.N., Trubilko, A.I., Zhiliba, A.I.: quant-ph//0011124 (2001)
14. Bose, S., Vedral, V., Knight, P.L.: Phys. Rev. A **57**, 822 (1998)
15. Karlsson, A., Bourennane, M.: Phys. Rev. A **58**, 4394 (1998)

Part VI
Extenics and Innovation Methods

Discipline Management System Reform in Extenics Perspective-Based on the Comparison Between the Old and New Discipline Catalogue

Yu-xuan Zhong and Hao Ren

Abstract From comparative analysis perspective between the old and the new discipline catalogue, the paper puts forward several hypotheses (research hypotheses and alternative hypotheses included) which need to be verified and establishes extension classification models based on extension data mining theories and methods in Extenics. Especially with models' application such as extensible set, elementary dependent function and evaluation information-element model, and through related verification, during which the promulgation and implementation of "new catalogue" is supposed as the corresponding extension transformation, the purpose of empirical study is achieved, also extension classification knowledge is gained. That is to say: Discipline management system and discipline catalogue are interrelated; The conduct of setting up and adjusting discipline and profession independently has never been implemented and the related power has never been expanded or enlarged, for which the issue of "new catalogue" of the year 2011 is an important and historic opportunity.

Keywords Discipline management system reform · Discipline catalogue · Extenics · Extension classification knowledge

1 Introduction

The importance of discipline management system reform is very obvious, as one part of higher education management system reform, which is especially involved in disciplinary construction and development for colleges or universities. Simultaneously in Zhong, Wu and Ren's opinion (2013) "Higher education management system is the key external motivation for disciplines' connotative development in research universities of science and engineering, of course which is most directly

Y. Zhong (✉) · H. Ren
School of Economics and Management, Tongji University, Shanghai 200092, China
e-mail: yxzhongzj@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_50

influenced by discipline management system" [1]. While Zhong explained (2011) "Generally in connotation-oriented development pattern, which is the key to quality, scientific discipline layout and rational structure are emphasized, outstanding characteristics and advantages of disciplines will be given full development" [2]. Disciplines in research universities of science and technology in China are widely developed in similarity and large scale nowadays, so the reform on management system in academic discipline appears more necessary and more important.

So far, many scholars have done research on discipline management system in China [3–9], which obviously has Chinese characteristics. Most of them are qualitative analysis on specific issues, such as general policy interpretation, theoretical analysis or cases study, etc. Thus the paper tries introducing theory and knowledge in Extenics-an original cross-discipline of China, which was founded by professor Cai W. in the 1980s, focusing on establishing empirical research models and analyzing discipline management system and its reform with quantitative analysis method and perspective.

2 Hypotheses and Extension Models

The academic degrees committee of state council and education ministry issued "the degree granting and talents cultivating discipline catalogue (2011 edition)" (abbreviation in "new catalogue") in March 2011. And degree-conferring units can independently set up secondary disciplines with the authorization of first-level disciplines at the moment. In fact, the project of setting up disciplines and professions independently in experimental units started officially in 2002 within the scope of first-level disciplines with doctorate authorization. In Bao's words (2006) "Discipline management system is the system of power distribution about disciplines setting up, degrees awarding and curriculums establishing in colleges or universities" [6]. Then there are two imperative questions proposed: Firstly, what is the substance of these two important reform acts of academic discipline management system-the policy of setting up disciplines and specialties respectively? Secondly, what is the difference between them?

Experimental work of the year 2002 started before the promulgation and implementation of "new catalogue", using the 1997 edition of discipline and profession catalogue for doctors' and masters' degrees awarding and postgraduate cultivating (abbreviation in "old catalogue"). There are discipline sorts, first-level and secondary disciplines with "old catalogue", which were adjusted into discipline sorts and first-level disciplines only in "new catalogue", adding one discipline category and twenty-one first level disciplines. So, what is the relationship between discipline management system reform and discipline catalogue? Has the action of setting up and adjusting disciplines independently been put into effect and enlarged in deed before or after the implementation of the new and the old catalogue?

Therefore, with problems oriented, several hypotheses (separately in research hypotheses and alternative hypotheses) in the following are put forward which need to be verified in this paper. The task of empirical research will be completed (e.g., confirmed or falsified), through analysis and verification with extension classification models' establishment and its application based on the theories and methods in extension data mining from the comparison perspective between the old and the new discipline catalogue.

Research hypothesis one: Discipline management system reform is unrelated with discipline catalogue;

Alternative hypothesis one: Discipline management system reform is related with discipline catalogue.

Research hypothesis two: The conduct of setting up and adjusting discipline and profession independently has been implemented and the corresponding power has been expanded or enlarged;

Alternative hypothesis two: The conduct of setting up and adjusting discipline and profession independently has never been implemented and the corresponding power has never been expanded or enlarged.

Frankly, extension data mining is the product of the combination of Extenics and data mining, which studies the mining of the knowledge relevant to the contradictory problem's solving or transformations from the database by using the theories and methods of Extenics [10–12]. Extension set theory, which is different from Cantor set and Fuzzy set, and extension logic are theoretically applied to data mining, adopting methods such as extension classification, extension conduction, using dependent functions as tools to complete quantitative calculation, to meet certain requirements and realize the combination of dynamic and static, qualitative and quantitative analysis effectively. From them, extension classification method is based on the extension transformation, including the extension classification based on the transformation of elements in the universe of discourse, dependent criterion and the universe of discourse. In the paper, element-based transformation will be taken for an example. All elements to be classified are denoted as universe of discourse U, in which evaluation characteristics are decided, and extension transformation as T for any elements, extensible set $\widetilde{E}(T)$ about U established.

Namely, suppose U is universe of discourse, u is any one element in U, k is a mapping of U to the real field $\mathfrak{R}$, $T = (T_U, T_k, T_u)$ is given transformation, in general call

$$\tilde{E}(T) = \{ (u, y, y') | u, \in U, y = k(u) \in \mathfrak{R}; \mathfrak{T}_{\mathfrak{u}} \mathfrak{u} \in \mathfrak{T}_{\mathfrak{U}} \mathfrak{U}, \mathfrak{y}' = \mathfrak{T}_{\mathfrak{k}} \mathfrak{k}(\mathfrak{T}_{\mathfrak{u}} \mathfrak{u}) \in \mathfrak{R} \}$$

an extension set on the universe of discourse U, $y = k(u)$ and $y' = T_k k(T_u u)$ are the dependent functions of $\widetilde{E}(T)$, wherein, T_U, T_k and T_u are transformations of respective universe of discourse U, dependent function k and element u.

When $T \neq e$, transformation T divides U into five fields, such as $E_+(T)$ (namely positive extensible field or positive qualitative change field), $E_-(T)$ (negative extensible field or negative qualitative change field), $E_+(T)$ (positive stable field or positive quantitative change field), E_- (T) (negative stable field or negative quantitative change field) and E_0 (T) (extension boundary). And when there is no transformation implemented, U can often be divided into three fields of E_+ (positive field), E_- (negative field), E_0 (zero boundary).

Then establish elementary dependent function $k(x)$ with midpoint of the interval not being the optimum considering real situation in the paper, especially evaluation characteristics based on time dimension (namely not discrete) and double intervals (standard and transition field) in general, as follows. Of course simple dependent function and discrete dependent function are available too.

$$k(x) = \frac{\rho(x, x_0, X_0)}{D(x, X_0, X)}. \tag{1}$$

In formula (1), x_0 is the optimal point, $X_0 = \langle a, b \rangle$ is desirable interval (e.g., standard positive field) and $X = \langle c, d \rangle$ is acceptable interval (e.g., positive field), where $x_0 \in \langle a, b \rangle$, $X_0 \subset X$ (generally X and X_0 have no common endpoints). $\rho(x, x_0, X_0)$ is side distance between x and interval X_0 from the side of point x_0; $D(x, X_0, X)$ is place value between point x and intervals X_0 and X. ρ_l (x, x_0, X_0) is called left side distance between x and interval X_0 about x_0, and right side distance ρ_r (x, x_0, X_0).

$$\rho l(x, x_0, X_0) = \begin{cases} a - x, x \le a \\ \frac{b - x_0}{a - x_0}(x - a),\ x \in \langle a, x_0 \rangle. \\ x - b, x \ge x_0 \end{cases} \tag{2}$$

$$\rho r(x, x_0, X_0) = \begin{cases} a - x, x \le x_0 \\ \frac{a - x_0}{b - x_0}(b - x\ ), x \in \langle x_0, b \rangle. \\ x - b, x \ge b \end{cases} \tag{3}$$

$$D(x,\ X_0,\ X) = \begin{cases} \rho(x, X) - \rho(x, X_0), \rho(x, X) \ne \rho(x, X_0), x \notin X_0 \\ \rho(x, X) - \rho(x, X_0) + a - b, \rho(x, X) \ne \rho(x, X_0), x \in X_0 \\ a - b, \rho(x, X) = \rho(x, X_0) \end{cases} \tag{4}$$

Among them, $\rho(x, X)$ is the extension distance between point x and interval X. And $\rho(x, X_0)$ is the distance between point x and interval X_0.

$$\rho(x, X) = \left| x - \frac{c + d}{2} \right| - \frac{d - c}{2}. \tag{5}$$

$$\rho(x, X_0) = \left| x - \frac{a + b}{2} \right| - \frac{b - a}{2}. \tag{6}$$

3 Empirical Analysis and Results Verified

3.1 Sample Selection and Data Sources

Revision task of "new catalogue" officially launched in 2009, while experimental project of setting up discipline and profession independently within the scope of first-level doctoral disciplines started in 2002. In the light of education ministry's official website information, relevant data from the year of 2002 to 2008 are taken as samples on time dimension (Table 1), only including the numbers of disciplines and specialties recorded, with eleven disciplines except military discipline, and the disciplines of philosophy, economics, law, education, literature, history joining in humanity and social science on the purpose of easy statistics and comparison, which referred to the "old catalogue".

3.2 Related Parameters and Indexes Illustration

Suppose information-element $I = (O, c, v)$, where O indicates original object, c and v stand for feature and its value, then $D = (I, d, u)$ is the single evaluation information-element for I in which d is the evaluation characteristic and u is evaluation value related. And $\{I\} = \{I_i \mid I_i = (O_i, C_i, V_i), i = 1, 2, \ldots, n\}$ is an information-element set used as a multi-dimensional evaluation information-element with multiple characteristics.

Therefore evaluation information-element D_i is built before the implementation of extension transformation φ considering real state of objects studied in the paper (as follows).

Table 1 Original data (time dimension)

Year[a]	Accumulative and recorded disciplines setting up independently of the year	Secondary disciplines for masters and doctors in catalogue
2002	653	367
2003	1023	367
2004	1253	367
2005	1459	367
2006	1665	367
2007	1761	367
2008	1855	367

[a]Data are the published data from education ministry's official website; Disciplines recorded which were cancelled each year have been considered and will never appear in "Accumulative and recorded disciplines setting up independently of the year", and all are counted only once; "Secondary disciplines for masters and doctors in catalogue" refers to "old catalogue"

$$D_i = [I_i, d, u_i].$$

$$I_i = \begin{bmatrix} Oi, & c6, & vi6 \\ & c7, & vi7 \\ & \vdots & \vdots \end{bmatrix}.$$

Take the ratio of secondary disciplines in "old catalogue", accumulative and recorded disciplines setting up independently of the year as evaluation characteristic values u_i for evaluation characteristic d, with original characteristics c_i and original characteristic values v_i. For evaluation characteristics, desirable interval X_0 = <a, b> and acceptable interval X = <c, d> should be decided according to maximum and minimum rule or method, namely X_0 = <0.2, 1>, X = <0.3, 1> (wherein, minimum 0.2, average 0.3, keep one digit after decimal point with rounding off method), K (x) has maximum when $x_0 = 1$.

Bao (2004) stated "Higher education institution has authority of setting up disciplines autonomously, which is one of academic discipline management system's characteristics in the United States. The autonomous right allows colleges and universities to get the substantial autonomy, rather than procedural autonomy. Meanwhile there is no direct intervention for the states to colleges and universities who can set up professional disciplines, which embodies government's action rule of 'do something preferred to and others never done' under the conditions of market economy" [8]. In fact, the United States' CIP-2000 to the largest extent records disciplines and specialties of advanced education, guaranteeing setting right of disciplines and specialties in colleges and universities, based on which there's theoretically assumption in this paper: if disciplines and specialties setting up autonomously that must be recorded are completely within the discipline catalogue, that is to say, when the evaluation characteristic values u_i are greater than or equal to number 1 (namely critical value $u_i = 1$), it means the conduct of setting up and adjusting disciplines and professions independently truly implemented and power expanded. In China since the "new catalogue" was promulgated in 2011, the government has been carrying out the task of statistics according to secondary disciplines with each university or college, which are compiled into the catalogue of secondary disciplines as a professional statistical catalogue. So the evaluation characteristic values u_i above are greater than or equal to 1, supposing the promulgation and implementation of "new catalogue" as transformation φ.

3.3 Results Analysis and Models Test

Calculate dependent function values through formula (1) in accordance with right side distance. And

$k(x) = \frac{\rho(x,1,X_0)}{D(x,X_0,X)}$, among which ρ (x, 1, X_0) = 0.3−x, $x \le 1$.

$$D(x, X_0, X) = \begin{cases} 0.1 & x = 0.293, 0.252, 0.220, 0.208, 0.198 \\ -0.6 & x = 0.562, 0.359 \\ & \vdots \end{cases}.$$

Then respective value before transformation φ are: k (u_1) = −0.437; k (u_2) = −0.098; k (u_3) = 0.07; k (u_4) = 0.48; k (u_5) = 0.8; k (u_6) = 0.92; k (u_7) = 1.02. Normalize dependent function values above in line with formula (2) as follows (process omitted here), and determine their extension types.

$$k_{ij} = \frac{k_j(u_{ij})}{\max\limits_{1 \le i \le 7} |k_j(u_{ij})|} (i = 1; j = 1, 2, \ldots, 7). \tag{7}$$

The value $k(u_i) = 1$ after implementing transformation φ, and the normalized value is 1 too, which belongs to E_+. Judge transformation types based on the correlation values before φ and after φ (Table 2).

Dependent d-value $\alpha_{ij} = k_j(u_{ij}') - k_j(u_{ij})$ and dependent p-value $\beta_{ij} = k_j(u_{ij}') \cdot k_j(u_{ij})$. Comprehensive dependent d-value $\alpha_i = K({}_{Ii}T_{Di}D_i) - K(D_i)$ and comprehensive p-value $\beta_i = K({}_{Ii}T_{Di}D_i) \cdot K(D_i)$. α_{min} and α_{max} express maximum and minimum of dependent d-value, β_{min} and β_{max} show the ones of dependent p-value (Table 3). There is no need to think about coefficient weights in the paper, so comprehensive dependent values are dependent values respectively.

And the following are classification knowledge according to Tables 2 and 3.

1. In Table 2, types of dependent values belong to either positive field E_+ or negative field E_-; Corresponding transformation types either have positive quantitative change E_+ (T) or positive qualitative change E_+ (T). And from Table 3, for evaluation characteristics d, when $0 \in [\alpha_{min}, \alpha_{max}]$, with extension transformation φ, some comprehensive dependent degrees of information-elements $\{I\}$ are increased, while some are reduced. When $\alpha_j > 0$, it means efficiency; If $\alpha_j = 0$, it stands for equivalent; when $0 \in [\beta_{min}, \beta_{max}]$, some

Table 2 Dependent values and classification

Names[a]	k_{1j}	Type	Transformation type
D_{2002}	−0.428	E_-	E_+ (T)
D_{2003}	−0.096	E_-	E_+ (T)
D_{2004}	0.069	E_+	E_+ (T)
D_{2005}	0.471	E_+	E_+ (T)
D_{2006}	0.784	E_+	E_+ (T)
D_{2007}	0.902	E_+	E_+ (T)
D_{2008}	1.000	E_+	E_+ (T)

[a]D2002–D2008 respectively stands for evaluation information in the year of 2002–2008; E_+ means positive field, E_- means negative field, E_+ (T) is positive qualitative change field, E_+ (T) is positive quantitative change field; Keep three digits after decimal point with round-off method

Table 3 Dependent d-value and p-value

Names[a]	Dependent d-value α_i	Dependent p-value β_i
D_{2002}	1.428	−0.428
D_{2003}	1.096	−0.096
D_{2004}	0.931	0.069
D_{2005}	0.529	0.471
D_{2006}	0.216	0.784
D_{2007}	0.098	0.902
D_{2008}	0	1.000
α_{min}	0	/
α_{max}	1.428	/
β_{min}	/	−0.428
β_{max}	/	1.000

[a]Sources of the original data are the same as those from Tables 1 and 2

information-elements make qualitative changes, while some make quantitative changes. When $\beta_j < 0$, it represents a qualitative change; If $\beta_j > 0$, it is a quantitative one. Therefore, if the second research hypothesis is true, namely autonomous power of setting up and adjusting disciplines and specialties has been executed and expanded, at least associated p-value (product value) β_j should show the property of homogeneity, all qualitative changes or all quantitative changes. But that's not the truth. So it's proved that research hypothesis two is false, while alternative hypothesis two is true: **The conduct of setting up and adjusting discipline and profession independently has never been implemented and the corresponding power has never been expanded or enlarged.**

2. In the knowledge equation, validity is measured by support and confidence which are usually expressed as ℓ = (Support, Confidence) = $\{(|E_-| + |E_0|)/|U|, |E_+(T)|/(|E_-| + |E_0|)\}$ for knowledge on positive qualitative change, and $\{|E_+|/|U|, |E_+(T)|/|E_+|\}$ for knowledge on positive quantitative change. The former indicates the support is the percentage that the sum $|E_-| + |E_0|$ of the objects in the negative field and zero boundary takes up in the number of all objects $|U|$ in the universe of discourse; the confidence is the percentage that the number $|E_+(T)|$ of the objects transformed from negative field or zero boundary into positive field takes up in the sum $|E_-| + |E_0|$ of the objects in negative field and zero boundary. The latter indicates the support is the percentage that the number $|E_+|$ of the objects in the positive field takes up in the number of all objects $|U|$ in the universe of discourse; the confidence is the percentage that the number of $|E_+(T)|$ of the objects that are still in the positive field after transformation takes up in the number $|E_+|$ of the objects in the positive field. The information-element set $\{I\}$ is right all objects $|U|$ in the paper, and T is the transformation φ. So in the paper the support $|E_-|/|U| = 2/7$ and confidence $|E_+(T)|/|E_-| = 2/2 = 1$ for positive qualitative change or $|E_+|/|U| = 5/7$ and $|E_+(T)|/|E_+| = 5/5 = 1$ for

positive quantitative change, considering $|E_0| = 0$. It is certain the results (value of $\ell > 0$) are all satisfactory.

3. From the data in results and their original data from Tables 2 and 3, before transformation φ carrying out evaluation characteristic value u_i of evaluation feature d is the ratio between secondary disciplines in "old catalogue" and accumulative and recorded disciplines setting up independently of the year. "New catalogue" only contains discipline sorts and first-level disciplines, and secondary discipline catalogue records the secondary disciplines setting up by colleges or universities themselves, making evaluation characteristic values greater than or equal to 1. It has theoretically proposed the assumption in the paper: if disciplines and specialties setting up independently which must be put on record are completely included in discipline catalogue, namely when evaluation characteristic value is greater than or equal to 1, it means the autonomous right of setting up and adjusting disciplines and professions is truly implemented and gets expansion. It's really true after the promulgation and implementation of "new catalogue", the government department in charge only checks degrees awarded of first-level disciplines, which is compared to the situation of "old catalogue", when our country had the secondary disciplines in management, and if degrees-conferring units wanted to adjust and set up secondary disciplines, they must get the nation's demonstration and approval. Of course this has also been confirmed by the first item above of the extension classification knowledge: although the experimental plan of the year 2002 for setting up disciplines and specialties independently started before "new catalogue" announced in 2011, the independent power has never been truly implemented and expanded. It can also see some scholars' involved papers, *Disciplinary management system reform from the perspective of transforming bridge in Extenics* for example [13]. Therefore, whether the autonomous right of setting up and adjusting disciplines and specialties is truly carried out or not is related to discipline catalogue tightly. **Finally it's concluded the reform of academic discipline management system is connected with discipline catalogue.**

4 Conclusion

With theories and methods of extension data mining in Extenics, especially through the establishment and application of extension classification models, extension classification knowledge has been gained and several hypotheses proposed have been verified. Before "new catalogue" was published in 2011, the power of independently setting up and adjusting disciplines and professions has never been truly implemented and expanded; Moreover, whether the independent power of setting up and adjusting disciplines and professions is really expanded or not is related to discipline catalogue closely, namely the academic discipline management system reform is concerned with discipline catalogue. Exactly to speak, expanding or enlarging the power of setting up and adjusting disciplines or professions independently is the main task of the discipline management system reform. The promulgation and implementation of "new catalogue" means the independent right of setting up and adjusting the discipline and specialty will get a real implementation and expansion, which is not only an important opportunity but also a great challenge for disciplinary system reform. It also has important historic significance for discipline development in connotation-oriented model or pattern, which undoubtedly ensures the external environment of discipline connotative development in research universities of science and engineering from the policy perspective.

Acknowledgements Thanks to the support by National Natural Science Foundation of China (No. 70872085, No. 71272048). All members from innovation research team at Tongji University Development Institute are gratefully and sincerely acknowledged. Tremendous thanks are also owned to Tongji University providing such a challenging platform especially for the fields of application research on Management, like the amazing and lovely research area of strategic alliance, innovation management, system reform and extension engineering, etc.

References

1. Zhong, Y.X., Wu, G.Y., Ren, H.: The analysis on key driving force of connotation-oriented disciplinary development on the basis of data mining in extenics-in the form of hypotheses testified with several research universities. In: Zhong, D., Hu, A. (eds.) Proceedings of the 2013 Asia-Pacific Computational Intelligence and Information Technology Conference, DEStech Publication, Inc., pp. 770–775 (2013)
2. Zhong, Y.X.: An empirical analysis on connotation-oriented disciplinary based on intrinsic motivation. Stud. Sci. Sci. **29**(2), 181–185 (2011)
3. Bao, R.: From 'planned supply' to 'market distribution': an evolution of the paradigm of college discipline and program management. J. Zhejiang Normal Univ. (Social Sciences) **32** (2), 1–5 (2007)
4. Liu, Z.T.: Further broadening autonomic power in running colleges and universities deepening reform of management system in disciplines and specialties. Res. High. Educ. **2**, 45–48 (2002)
5. Zhang, Z.G., Xiang, L.R.: The systemic research on disciplines catalogue in American higher education. Acad. Degrees Graduate Educ. 70–77 (2008)
6. Bao, R.: A case in India and the analysis on type of disciplines management system. Higher Educ. Chem. Eng. **6**, 15–18 (2006)
7. Wang, L.H., Gu, J.M.: The management systems and operation mechanisms of academic program in the U.S. higher education systems. Res. Higher Educ. Eng. **3**, 90–94 (2009)
8. Bao, R.: The features of the U.S. 'classification of instructional programs' and its implications. Comp. Educ. Rev. **4**, 1–5 (2004)
9. Wang, J.H.: Management of discipline and specialty: a new field of higher education research. J. Qingdao Univ. Sci. Technol. (Social Sciences) **23**(1), 58–62 (2007)
10. Cai, W., Yang, C.Y.: Extensible Set and Extension Data Mining. Science Press, Peking (2008)
11. Yang, C.Y., Cai, W.: Extension Engineering. Science Press, Peking (2007)
12. Yang, C.Y., Cai, W.: Extenics Theory, Method and Application. Science press, Peking (2013)
13. Zhong, Y.X., Fang, Y.M., Ren, H.: Disciplinary management system reform from the perspective of transforming bridge in Extenics. Jiangsu Higher Educ. **6**, 48–49 (2012)

A Logical Representation of Extension Transformation

Zhi-bin Chen, Xi-ping Jia and Zheng-hong Xiao

Abstract In this paper extension transformation is considered as a kind of relation between the object accepted transformation and the objects as transformation result. For studying the logical representation of extension transformation, the theory of extension logic is developed further, including its syntax and semantics. The logical definitions about extension transformation, basic operations of extension transformation and conductive transformation are given with some examples.

Keywords Extenics · Extension logic · Extension transformation · Conductive transformation

1 Introduction

In Extenics, extension transformation is one of the key concepts and it is also the core tool for solving contradictory problems. According to [1], the general definition of extension transformation is given briefly as follow: "Suppose Γ_0 is any object of basic-element, compound-element, criterion, and the universe of discourse. The transformation from Γ_0 to the object of the same class Γ or multiple objects in the same class $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ is referred to as extension transformation of object Γ_0, denoted as $T\Gamma_0 = \Gamma$ or $T\Gamma_0 = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_n\}$."

$T\Gamma_0 = \Gamma$ can be denoted as $T = (\Gamma_0, \Gamma)$ [1], which better expresses that extension transformation can be comprehended as not only the unary operator, but also a kind of relation between the object accepted transformation and the objects as transformation result. From this, the fundamental question that needs to be answered is how Γ_0 relate to Γ in $T\Gamma_0 = \Gamma$. In other words, given the expression $T = (\Gamma_0, \Gamma)$, how can we decide whether T is an extension transformation or not?

Z. Chen (✉) · X. Jia · Z. Xiao
School of Computer Science, Guangdong Polytechnic Normal University,
510665 Guangzhou, China
e-mail: cyblink@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_51

For answering the above question, the logical definition of extension transformation is need, that is, we need to study the problem of logical representation of extension transformation. As a result in this paper we develop the theory of extension logic further, introducing the syntax and the semantics of our proposed logic. After this, we give the logical definitions about extension transformation, basic operations of extension transformation and conductive transformation with some examples.

2 Syntax and Semantics of Extension Logic

2.1 Syntax

We introduce a language **L** for studying the syntax of extension logic.

Definition 1 (*Alphabet*) An alphabet for **L** consists of the following symbols:

(i) *Sort* is a countable set of sort symbols. We use s_o, s_c, s_v to denote three most general sorts about object, characteristic and measure respectively, where $s_o, s_c, s_v \in$ *Sort*.
(ii) *Fun* is a countable set of function symbols. We use Fun_n to denote the set of n-ary function symbols, $Fun = \bigcup_{n\geq 0} Fun_n$. The symbols in Fun_0 are called constants.
(iii) *Bas* is a countable set of basic-element symbols. We use Bas_n to denote the set of n-dimensional basic-element symbols, $Bas = \bigcup_{n\geq 1} Bas_n$.
(iv) *Var* is a countably infinite set of variables. We use Var_s to denote the set of variables of sort s, $Var = \bigcup_{s\in sort} Var_s$.
(v) $\neg$, $\vee$, $\wedge$, $\rightarrow$, $\mapsto$ are the logical connectives.
(vi) "[", "]", "(", ")", ",", "=" are the technical symbols.

Definition 2 (*Declaration*) A declaration over *Sort* ∪ *Fun* ∪ *Bas* for **L** is an ordered triple $Decl = (D_s, D_F, D_B)$ such that

(i) D_S is a set of subsort declarations of the form $\langle s', s\rangle$, where $s', s \in Sort$. For any sort s', $\langle s', s\rangle \in D_S$, where $s \in \{s_o, s_c, s_v\}$.
(ii) D_F is a set of function declarations of the form $f : \langle s_1, \ldots, s_n, s\rangle$, where $f \in Fun_n (n \geq 0)$, $s_1, \ldots, s_n, s \in Sort$.
(iii) D_B is a set of basic-element declarations of the form $B : \langle s^O, s^C_1, \ldots, s^C_n, s^V_1, \ldots, s^V_n\rangle$, where $s^O, s^C_1, \ldots, s^C_n, s^V_1, \ldots, s^V_n \in Sort$.

Definition 3 (*Signature*) A signature for language **L** is an ordered quadruple $\Sigma = (Sort, Fun, Bas, Decl)$, where *Sort* is the set of all sort symbols, *Fun* is the set of all function symbols, *Bas* is the set of all basic-element symbols, and *Decl* is the declaration over *Sort* ∪ *Fun* ∪ *Bas*.

Definition 4 (Σ-*TERM*) Let Σ be a signature. Given sort $s \in Sort$, the set Σ-TERM(s) of Σ-terms of sort s is defined by the following rules:

(i) $Var_s \subseteq \Sigma$-TERM(s).
(ii) For any constants con, if $con : \langle s \rangle \in D_F$, then $con \in \Sigma$-TERM(s).
(iii) If $t_1, \ldots, t_n$ are Σ-terms of sorts $s_1, \ldots, s_n$, then $f(t_1, \ldots, t_n) \in \Sigma$-TERM($s$), where $f \in Fun_n$ and $f : \langle s_1, \ldots, s_n, s \rangle \in D_F$.
(iv) If t is the Σ-term of sort s', then $t \in \Sigma$-TERM(s) where $\langle s', s \rangle \in D_S$.

Definition 5 (Σ-*BASIC*) Let Σ be a signature. Σ-BASIC is the set of Σ-basic-elements. A Σ-basic-elements is an expression of the form

$$B = \begin{bmatrix} O, & c_1, & v_1 \\ & c_2, & v_2 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix}$$

where $B \in Bas$ is n-dimensional basic-element symbol, $B : \langle s^O, s_1^C, \ldots, s_n^C, s_1^V, \ldots, s_n^V \rangle \in D_B$, $O, c_1, \ldots, c_n, v_1, \ldots, v_n$ are Σ-terms of sorts $s^O, s_1^C, \ldots, s_n^C, s_1^V, \ldots, s_n^V$.

Definition 6 (Σ-*FORM*) Let Σ be a signature. The set Σ-FORM of Σ-Formulas is defined by the following rules:

(i) Σ-BASIC $\subseteq \Sigma$-FORM.
(ii) If $B_1, B_2 \in \Sigma$-BASIC, then $(B_1 \mapsto B_2) \in \Sigma$-FORM.
(iii) If $\alpha \in \Sigma$-FORM, then $(\neg\alpha) \in \Sigma$-FORM.
(iv) If $\alpha_1, \alpha_2 \in \Sigma$-Form, $(\alpha_1 \vee \alpha_2)$, $(\alpha_1 \wedge \alpha_2)$, $(\alpha_1 \rightarrow \alpha_2) \in \Sigma$-Form.

Note that for simplicity sometimes we denote the Σ-basic-element only by its basic-element symbol. We omit some parentheses of Σ-formula as usual. The Σ-formula without variable is said to be Σ-sentence.

2.2 Semantics

Definition 7 (*Situation*) Let U_O, U_C, U_V be the nonempty sets of individuals of object, characteristic and measure. Let $R = U_O \times U_C \times U_V$. Situation sit is a nonempty set where $sit \subseteq R$. sit' is a subsituation of situation sit, where $sit' \subseteq sit$. $H(sit)$ denotes all subsituations of sit, that is, $H(sit) = \{sit' | sit' \subseteq sit, sit' \text{is nonempty set}\}$.

Definition 8 (*World*) The set W of worlds is a set of situations such that $\bigcup_{i=1}^{n} sit_i = R$ and $sit_1, \ldots, sit_n$ differ in some respects. One of the worlds in W is said to be real world w_r, the other worlds are called possible worlds. We call sit is a situation of w if $sit \subseteq w$.

Definition 9 (*Transformation mapping*) Let $\mathrm{R} = \mathrm{U}_O \times \mathrm{U}_C \times \mathrm{U}_V$, for any situation $sit \subseteq \mathrm{R}$, transformation mapping tr_{sit} of sit is defined by $tr_{sit} : H(sit) \to P(H(\mathrm{R}))$, where $H(sit)$ and $H(\mathrm{R})$ are the set of subsituations of sit and R, $P(H(\mathrm{R}))$ is the power set of $H(\mathrm{R})$. Let $\mathrm{T} = \bigcup_{sit \in H(\mathrm{R})} tr_{sit}$.

Definition 10 (*Extension space*) An extension space is an ordered quintuple $\Omega = (\mathrm{U}_O, \mathrm{U}_C, \mathrm{U}_V, \mathrm{R}, \mathrm{T})$.

Definition 11 (*Σ-Interpretation*) Let $\Sigma = (Sort, Fun, Bas, Decl)$ be a signature, a Σ-interpretation M in language **L** is a pair (Ω, I), where Ω is an extension space, $\Omega = (\mathrm{U}_O, \mathrm{U}_C, \mathrm{U}_V, \mathrm{R}, \mathrm{T})$, I is a function such that:

(i) For $s \in Sort$, $I(s) = \mathrm{U}_s$, where U_s is the nonempty set of individuals of sort s. In particular, $I(s_o) = \mathrm{U}_O$, $I(s_c) = \mathrm{U}_C$ and $I(s_v) = \mathrm{U}_V$.
(ii) For $\langle s', s \rangle \in D_S$, $I(s') \subseteq I(s)$.
(iii) For $f \in Fun_n$ with $f : \langle s_1, \dots, s_n, s \rangle \in D_F$, $I(f) : I(s_1) \times \cdots \times I(s_n) \to I(s)$.
(iv) For $B \in Bas_n$ with $B : \langle s^O, s_1^C, \dots, s_n^C, s_1^V, \dots, s_n^V \rangle \in D_B$,
$$I(B) \subseteq \bigcup_{i=1}^{n} \{I(s^O) \times I(s_i^C) \times I(s_i^V)\}\}$$

Definition 12 (*Evaluation*) Let $M = (\Omega, I)$ be a Σ-interpretation. A variable assignment is a function $\mu : Var \to \mathrm{U}_O \cup \mathrm{U}_C \cup \mathrm{U}_V$ such that $\mu(x) \in I(s)$ for all variables $x \in Var_s$. Let t be a Σ-term. $\bar{\mu}(t)$ is called the evaluation function of Σ-term. $\bar{\mu}(t)$ is recursively defined by the following rules:

(i) For $x \in Var_s$, $\bar{\mu}(x) = \mu(x)$.
(ii) For constant con, $\bar{\mu}(con) = I(con)$.
(iii) For Σ-term $f(t_1, \dots, t_n)$, $\bar{\mu}(f(t_1, \dots, t_n)) = I(f)(\bar{\mu}(t_1), \dots, \bar{\mu}(t_n))$.

Definition 13 (*Satisfiability*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. Given situation sit, according to the following rules we say that the Σ-formula α is satisfied in sit under M, written $(\alpha)^{sit,M,\mu} = 1$, otherwise $(\alpha)^{sit,M,\mu} = 0$:

(i) $\left(B = \begin{bmatrix} O, & c_1, & v_1 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix} \right)^{sit,M,\mu} = 1$ if and only if $\bigcup_{i=1}^{n} \{\langle \bar{\mu}(O), \bar{\mu}(c_i), \bar{\mu}(v_i) \rangle\} \subseteq I(B)$ and $I(B) \subseteq sit$.
(ii) $(B_1 \mapsto B_2)^{sit,M,\mu} = 1$ if and only if $(B_1)^{sit,M,\mu} = 1$ and there exists $sit' \in tr_{sit}(I(B_1))$ such that $(B_2)^{sit',M,\mu} = 1$, where tr_{sit} is the transformation mapping of sit.
(iii) $(\neg\alpha)^{sit,M,\mu} = 1$ if and only if $(\alpha)^{sit,M,\mu} = 0$.
(iv) $(\alpha_1 \vee \alpha_2)^{sit,M,\mu} = 1$ if and only if $(\alpha_1)^{sit,M,\mu} = 1$ or $(\alpha_2)^{sit,M,\mu} = 1$.
(v) $(\alpha_1 \wedge \alpha_2)^{sit,M,\mu} = 1$ if and only if $(\alpha_1)^{sit,M,\mu} = 1$ and $(\alpha_2)^{sit,M,\mu} = 1$.
(vi) $(\alpha_1 \to \alpha_2)^{sit,M,\mu} = 1$ if and only if $(\alpha_1)^{sit,M,\mu} = 0$ or $(\alpha_2)^{sit,M,\mu} = 1$.

We write $(\alpha)^{sit,M} = 1$ when α is a Σ-sentence and satisfied in sit under M.

Proposition 1 *Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. Given Σ-basic-element B, situation sit and sit', $sit \subseteq sit'$. If $(B)^{sit,M,\mu} = 1$, then $(B)^{sit',M,\mu} = 1$.*

Proof Let $\left(B = \begin{bmatrix} O, & c_1, & v_1 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix} \right)^{sit,M,\mu} = 1.$

Then $\bigcup_{i=1}^{n} \{\langle \bar{\mu}(O), \bar{\mu}(c_i), \bar{\mu}(v_i) \rangle\} \subseteq I(B)$ and $I(B) \subseteq sit$. Because $sit \subseteq sit'$, we have $I(B) \subseteq sit'$. According to definition of satisfiability of Σ-formula, we see that $(B)^{sit',M,\mu} = 1$.

Definition 14 Let Ψ be the set of Σ-formulas. Given situation *sit*, we say that Ψ is satisfiable in sit if and only if there exists a Σ-interpretation $M = (\Omega, I)$ and variable assignment μ, such that for every $\alpha \in \Psi$, $(\alpha)^{sit,M,\mu} = 1$.

Definition 15 (*Logical consequence*) Let $\Psi \subseteq \Sigma$-FORM, $\alpha \in \Sigma$-FORM. We write $\Psi \vDash \alpha$ and say that α is a logical consequence of Ψ if and only if $\Psi \cup \{\neg\alpha\}$ is unsatisfiable in any situation.

We write $\Psi \nvDash \alpha$ if $\Psi \vDash \alpha$ is not true.

Proposition 2 *$(B_1 \mapsto B_2) \vDash B_1$, where B_1, B_2 are Σ-basic-elements.*

Proof Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. For any situation *sit*, if $(B_1 \mapsto B_2)^{sit,M,\mu} = 1$, we have ${B_1}^{sit,M,\mu} = 1$, then $(\neg B_1)^{sit,M,\mu} = 0$, therefore $\{B_1 \mapsto B_2, \neg B_1\}$ is unsatisfiable in *sit*, that is, $(B_1 \mapsto B_2) \vDash B_1$.

There are similar propositions such as $B_1 \mapsto B_2 \nvDash B_2, B_1, B_2 \nvDash (B_1 \mapsto B_2), (B_1 \mapsto B_2) \nvDash (B_2 \mapsto B_1)$, etc. We omit the proofs of these propositions.

3 Logical Representation

3.1 Definitions

Based on the general definition of extension transformation in [1], we introduce the definition of Σ-extension transformation.

Definition 16 (Σ-*Extension transformation*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. Given situation *sit*, we say that Σ-formula $B_0 \mapsto B$ is a Σ-extension transformation in *sit*, if and only if $(B_0 \mapsto B)^{sit,M,\mu} = 1$.

We use the purposed theory of extension logic to define Σ-extension transformation. Σ-formula $B_0 \mapsto B$ is the logical representation of the expression $T\Gamma_0 = \Gamma$, where Γ_0, Γ are the basic-elements represented by Σ-basic-elements B_0, B. The definition of Σ-extension transformation shows that, we consider the expression

$T\Gamma_0 = \Gamma$ is an extension transformation only when the corresponding Σ-formula $B_0 \mapsto B$ is satisfied in given situation.

Likewise, the extension transformation such as $T\Gamma_0 = \{\Gamma_1, \Gamma_2, \dots, \Gamma_n\}$ can be represented through a set of Σ-extension transformations $\{B_0 \mapsto B_1, B_0 \mapsto B_2, \dots, B_0 \mapsto B_n\}$. Note that besides basic-element, the object of extension transformation can be compound-element, criterion and universe of discourse. Because we can represent the compound-element through a set of basic-elements, meanwhile as the objects to be analyzed both criterion and universe of discourse can also be represented by basic-elements, therefore the proposed theory can also be applied for representing the extension transformations about compound-element, criterion and universe of discourse.

We introduce the logical representation of basic operations of extension transformation and conductive transformation based on the corresponding definitions in [1].

Definition 17 (*Target situation set*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. Let $B_0 \mapsto B$ be a Σ-extension transformation in situation *sit*. We call that *sit* is the source situation of $B_0 \mapsto B$ and sit' is the target situation of $B_0 \mapsto B$, where $sit' \in tr_{sit}(I(B_0))$ and $(B)^{sit',M,\mu} = 1$. We use $TS_{sit}(B_0 \mapsto B)$ to denote the set of target situations of $B_0 \mapsto B$.

Definition 18 (*Basic operations of* Σ*-extension transformation*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. The definitions of basic operations of Σ-extension transformation are given as follow:

(i) We use $\alpha_1, \dots, \alpha_n$ to denote the Σ-extension transformations $B_1 \mapsto B_2, \dots, B_{n-1} \mapsto B_n$ in situation $sit_1, \dots, sit_n$, and use β to denote the Σ-extension transformation $B_1 \mapsto B_n$ in situation sit_1. We say that β is the Σ-PRODUCT transformation of $\alpha_1, \dots, \alpha_n$ in sit_1 if and only if $sit_{i+1} \in TS_{sit_i}(\alpha_i)(i = 1, \dots, n-1)$ and $TS_{sit_n}(\alpha_n) = TS_{sit_1}(\beta)$.

(ii) We use $\alpha_1, \dots, \alpha_n$ and β to denote the Σ-extension transformations in situation *sit*. We say that β is the Σ-AND transformation of $\alpha_1, \dots, \alpha_n$ in *sit* if and only if $(\bigcap_{i=1}^{n} TS_{sit}(\alpha_i)) = TS_{sit}(\beta)$.

(iii) We use $\alpha_1, \dots, \alpha_n$ and β to denote the Σ-extension transformations in situation *sit*. We say that β is the Σ-OR transformation of $\alpha_1, \dots, \alpha_n$ in *sit* if and only if $(\bigcup_{i=1}^{n} TS_{sit}(\alpha_i)) = TS_{sit}(\beta)$.

(iv) We use α, β to denote the Σ-extension transformations $B_0 \mapsto B_1$, $B_1 \mapsto B_0$ in situation sit_1, sit_2. We say that β is the Σ-INVERSE transformation of α in sit_2 and α is the Σ-INVERSE transformation of β in sit_1 if and only if $sit_2 \in TS_{sit_1}(\alpha)$ and $sit_1 \in TS_{sit_2}(\beta)$.

Definition 19 (Σ-*First order conductive transformation*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. We use α, β to denote the Σ-extension transformations in situation *sit*. We say that β is the Σ-first order conductive transformation of α and α is active transformation of β in *sit* if and only if $TS_{sit}(\alpha) \subseteq TS_{sit}(\beta)$.

Given Σ-extension transformations α, β, we use $\alpha \Rightarrow \beta$ to denote that β is the Σ-first order conductive transformation of α.

Definition 20 (Σ-*n-times conductive transformation*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. We use α and $\beta_1, \ldots, \beta_n$ to denote the Σ-extension transformations in situation *sit*. We say that β_n is the Σ-n-times conductive transformation of α in *sit* if and only if $\alpha \Rightarrow \beta_1$, $\beta_1 \Rightarrow \beta_2$, $\ldots, \beta_{n-1} \Rightarrow \beta_n$.

Definition 21 (Σ-*m-order transformation and m-order conductive transformation*) Let $M = (\Omega, I)$ be a Σ-interpretation and let μ be the variable assignment. Let m be integer, $m > 1$. Given situation *sit*, $\alpha^{(m-1)}$ denotes the Σ-extension transformations $B_0^{(m-1)} \mapsto B_0^{(m)}$ and $\beta_i^{(m-1)}(i = 1, \ldots, n)$ denote the Σ-extension transformations $B_i^{(m-1)} \mapsto B_i^{(m)}$, such that $\beta_n^{(m-1)}$ is Σ-n-times conductive transformation of $\alpha^{(m-1)}$ in *sit*. Let $\alpha^{(m)}$ denotes the Σ-extension transformation $B_0^{(m)} \mapsto B_0^{(m+1)}$. We say that $\alpha^{(m)}$ is Σ-m-order transformation of $B_0^{(1)}$ in *sit* if and only if $\beta_n^{(m-1)} \Rightarrow \alpha^{(m)}$, and $\beta_n^{(m)}$ is the Σ-m-order n-times conductive transformation of $B_n^{(1)}$ in *sit* if and only if $\alpha^{(m)} \Rightarrow \beta_1^{(m)}, \beta_1^{(m)} \Rightarrow \beta_2^{(m)}, \ldots, \beta_{n-1}^{(m)} \Rightarrow \beta_n^{(m)}$.

3.2 Examples

Example 1 Representation of extension transformation in real world.

We take the extension transformation in the classical contradictory problem "Cao Chong weighed the elephant" [2] as example. Given Σ-basic-elements

$$B_1 = \begin{bmatrix} \text{elephant } A, & \text{weight}, & x \\ & \text{divisibility}, & 0 \end{bmatrix}, B_2 = \begin{bmatrix} \text{stone } B, & \text{weight}, & x \\ & \text{divisibility}, & 1 \end{bmatrix},$$

where variable $x \in Var_{s_v}$.

Given extension space Ω, let w_r be the real world. Let situation $sit_1, sit_2 \subset w_r$, where

$$sit_1 = \{\langle\bar{\mu}(\text{elephant } A), \bar{\mu}(\text{weight}), \bar{\mu}(500\,\text{kg})\rangle, \ldots,$$
$$\langle\bar{\mu}(\text{elephant } A), \bar{\mu}(\text{weight}), \bar{\mu}(10000\,\text{kg})\rangle,$$
$$\langle\bar{\mu}(\text{elephant } A), \bar{\mu}(\text{divisibility}), \bar{\mu}(0)\rangle\}$$

$$sit_2 = \{\langle\bar{\mu}(\text{stone } B), \bar{\mu}(\text{weight}), \bar{\mu}(500\,\text{kg})\rangle, \ldots,$$
$$\langle\bar{\mu}(\text{stone } B), \bar{\mu}(\text{weight}), \bar{\mu}(10000\,\text{kg})\rangle,$$
$$\langle\bar{\mu}(\text{stone } B), \bar{\mu}(\text{divisibility}), \bar{\mu}(1)\rangle\}$$

Let $M = (\Omega, I)$ be a Σ-interpretation such that $I(B_1) = sit_1$, $I(B_2) = sit_2$. Given the transformation mapping tr_{w_r} such that $sit_2 \in tr_{w_r}(sit_1)$. Given the variable assignment μ such that $\mu(x) \in \{\bar{\mu}(500\,\text{kg}), \ldots, \bar{\mu}(10000\,\text{kg})\}$. Because $sit_1 \subset w_r$, with proposition 1 we have $(B_1)^{w_r,M,\mu} = 1$. Obviously $(B_2)^{sit_2,M,\mu} = 1$, then $(B_1 \mapsto B_2)^{w_r,M,\mu} = 1$. As a result we call that $B_1 \mapsto B_2$ is a Σ-extension transformation in real world.

Example 2 Possible world and extension transformation.

Given Σ-basic-elements

$B_1 = \left[\text{person } A, \text{ height}, \ 1.70\,\text{m}\right], B_2 = \left[\text{person } A, \text{ height}, \ 1.75\,\text{m}\right]$.

Given extension space Ω, let w_r be the real world. Let situation $sit_1, sit_2 \subset w_r$, where

$sit_1 = \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.70\,\text{m})\rangle\}$,

$sit_2 = \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.75\,\text{m})\rangle\}$,

$\{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(0.0\,\text{m})\rangle, \ldots, \langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.8\,\text{m})\rangle\} \subset w_r$.

Let $M = (\Omega, I)$ be a Σ-interpretation such that $I(B_1) = sit_1, I(B_2) = sit_2$. Given the transformation mapping tr_{w_r} such that

$tr_{w_r}(I(B_1)) = \{\{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.70\,\text{m})\rangle\}, \ldots, \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.8\,\text{m})\rangle\}\}$

$tr_{w_r}(I(B_2)) = \{\{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.75\,\text{m})\rangle\}, \ldots, \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.8\,\text{m})\rangle\}\}$

Evidently $(B_1)^{sit_1,M} = 1$ and $(B_2)^{sit_2,M} = 1$, thus we have $(B_1)^{w_r,M} = 1$ and $(B_2)^{w_r,M} = 1$. Because $sit_2 \in tr_{w_r}(I(B_1))$, therefore we have $(B_1 \mapsto B_2)^{w_r,M} = 1$. On the other hand, because for every $sit' \in tr_{w_r}(I(B_2))$, we see that $(B_1)^{sit',M} = 0$, thus $(B_2 \mapsto B_1)^{w_r,M} = 0$.

The intuitive meaning of $B_1 \mapsto B_2$ can be the height increase of person A. Note that according to our definition of situation and world, real world w_r that we discuss here is a set of relevant situations about the object "person A", including all possibility of his height, thus both B_1 and B_2 can be satisfied in w_r, however only B_1 can be satisfied in sit_1.

Let $B_3 = \left[\text{person } A, \text{ height}, \ 3\,\text{m}\right]$. Given a possible world w where

$\{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(0.0\,\text{m})\rangle, \ldots, \langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(3\,\text{m})\rangle\} \subseteq w$.

But we see that

$\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(1.81\,\text{m})\rangle, \ldots, \langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(3\,\text{m})\rangle \notin w_r$

Given a situation $sit_3 = \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(3\,\text{m})\rangle\}$ and the transformation mapping tr_w such that

$tr_w(sit_1) =$

$\{\{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(0.0\,\text{m})\rangle\}, \ldots, \{\langle\bar{\mu}(\text{person } A), \bar{\mu}(\text{height}), \bar{\mu}(3\,\text{m})\rangle\}\}$,

and $tr_w(sit_2) = tr_w(sit_3) = tr_w(sit_1)$.

Let $I(B_3) = sit_3$. Because for every situation $sit' \in tr_{w_r}(I(B_1))$, $(B_3)^{sit',M} = 0$, therefore $(B_1 \mapsto B_3)^{w_r,M} = 0$. On the other hand, because $sit_3 \in tr_w(I(B_1))$ and $(B_1)^{w,M} = 1$, $(B_3)^{sit_3,M} = 1$, therefore $(B_1 \mapsto B_3)^{w,M} = 1$, that is, $B_1 \mapsto B_3$ is a Σ-extension transformation in w.

Example 3 Representation of conductive transformation.

Given Σ-basic-elements

$B_1 = \left[\text{product } A, \text{ price}, \ 2.0 \text{ dollars}\right], B_2 = \left[\text{product } A, \text{ price}, \ 2.5 \text{ dollars}\right]$,

$B_1{}' = \left[\text{product } B, \text{ price}, \ 4.0 \text{ dollars}\right], B_2{}' = \left[\text{product } B, \text{ price}, \ 5.0 \text{ dollars}\right]$.

Given Σ-interpretation $M = (\Omega, I)$ and situation

$sit =$
$\{\langle\bar{\mu}(\text{product } A), \bar{\mu}(\text{price}), \bar{\mu}(2.0 \text{ dollars})\rangle,$
$\langle\bar{\mu}(\text{product } B), \bar{\mu}(\text{price}), \bar{\mu}(4.0 \text{ dollars})\rangle\}$
such that $B_1 \mapsto B_2$ and $B_1{}' \mapsto B_2{}'$ are Σ-extension transformations in *sit*. If we have
$TS_{sit}(B_1 \mapsto B_2) =$
$\{\{\langle\bar{\mu}(\text{product } A), \bar{\mu}(\text{price}), \bar{\mu}(2.5 \text{ dollars})\rangle,$
$\langle\bar{\mu}(\text{product } B), \bar{\mu}(\text{price}), \bar{\mu}(5.0 \text{ dollars})\rangle\}\}$
$TS_{sit}(B_1{}' \mapsto B_2{}') = \{$
$\{\langle\bar{\mu}(\text{product } A), \bar{\mu}(\text{price}), \bar{\mu}(2.0 \text{ dollars})\rangle,$
$\langle\bar{\mu}(\text{product } B), \bar{\mu}(\text{price}), \bar{\mu}(5.0 \text{ dollars})\rangle\},$
$\{\langle\bar{\mu}(\text{product } A), \bar{\mu}(\text{price}), \bar{\mu}(2.5 \text{ dollars})\rangle,$
$\langle\bar{\mu}(\text{product } B), \bar{\mu}(\text{price}), \bar{\mu}(5.0 \text{ dollars})\rangle\},$
$\{\langle\bar{\mu}(\text{product } A), \bar{\mu}(\text{price}), \bar{\mu}(3.0 \text{ dollars})\rangle,$
$\langle\bar{\mu}(\text{product } B), \bar{\mu}(\text{price}), \bar{\mu}(5.0 \text{ dollars})\rangle\}\}$
because $TS_{sit}(B_1 \mapsto B_2) \subseteq TS_{sit}(B_1{}' \mapsto B_2{}')$, we can say that $B_1{}' \mapsto B_2{}'$ is the Σ-first order conductive transformation of $B_1 \mapsto B_2$ in *sit*.

4 Conclusion

From the viewpoint of epistemology, extension transformation, basic operations of extension transformation and conductive transformation represent the knowledge about possibility of objects change. In this paper we present some preliminary research results of extension logic and using the proposed theory for studying the logical representation of extension transformation. We introduce the concept of situation which is also applied in other relevant research fields, such as situation theory [3], reasoning about action [4], analogy and metaphor [5, 6], etc. On the basis of these theories, we give our definition of situation and represent the changes of objects as the mappings of situations.

Beginning with the first research book published in 2003 [2], the continuing studies of extension logic and the related topics about knowledge representation and reasoning can be found in [7–13]. On the whole for establishing the complete theoretical framework of extension logic much further research is needed. In the future we will focus on the study of logical extension reasoning based on extension transformation and the application of logical extension reasoning for supporting creativity activity such as design, contradictory problem solving, etc. Moreover, logical representation and reasoning of the relations between extension transformations is another important point in our future research.

Acknowledgments This work is supported by the Scientific Research Project of Education Department of Guangdong Province (No. 2013KJCX0117).

References

1. Yang, C.Y., Cai, W.: Extenics. Science Press, Beijing (2014)
2. Cai, W., Yang, C.Y., He, B.: Preliminary Extension Logic. Science Press, Beijing (2003)
3. Seligman, J.M, Moss, L.: Situation Theory. In: van Benthem, ter Meulen A (eds.) Handbook of Logic and Language, pp. 253–328, 2nd edn. Elsevier, Amsterdam (2010)
4. Raymond, R.: Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems. MIT Press, Cambridge (2001)
5. Keith, J.H.: Analogy. In: Keith, J.H., Robert, G.M. (eds.) The Cambridge Handbook of Thinking and Reasoning, pp. 117–142 . Cambridge University Press, Cambridge (2005)
6. Steinhart, E.C.: The Logic of Metaphor: Analogous Parts of Possible Worlds. Kluwer Academic Publishers, Dordrecht (2001)
7. Yang, C.Y., Wang, G.H., Li, Y., Cai, W.: Study on knowledge reasoning based on extended formulas. In: Li D.L., Wang, B.J. (eds.) Artificial Intelligence Applications and Innovations. Springer, New York (2005)
8. Yang, C.Y., Cai, W.: Knowledge representations based on extension rules. In: Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, pp. 25–27 (2008)
9. Chen, W.W., Yang, C.Y., Huang, J.C.: Extension knowledge and extension knowledge reasoning. J. Harbin Inst. Technol. **38**(7), 1094–1096 (2006)
10. Sun, H.A.: Extension propositions and extension consequences. J. Harbin Inst. Technol. **38**(7), 1101–1103 (2006)
11. Jiang, Z.W., Yao, M., Shen, B., Yi, W.S.: Extension knowledge spaces and extension knowledge grid model (I): extension knowledge spaces. J. Circuits Syst. **11**(4), 131–135 (2006)
12. Jiang, Z.W., Yao, M., Shen, B., Yi, W.S.: Extension knowledge spaces and extension knowledge grid model (II): extension knowledge grid. J. Circuits Syst. **11**(4), 136–141 (2006)
13. Chen, Z.B.: Study on semantics of extension logic. In: Cai, W. et al. (eds.) Extenics and Innovation Methods, pp. 43–48. CRC Press, Leiden (2013)

A Novel Establishing Method for 1-Dimensional Elementary Dependent Function by Standard Interval Transformation

Long Tang and Chunyan Yang

Abstract Decision-making commonly demands reliable evaluating conclusions of the candidate schemes as important references. Superiority evaluation is an effective evaluating method, which employs elementary dependent function to describe the degree of a candidate scheme satisfying the requirement and has been successfully applied to various fields. However, traditional elementary dependent function supplies only one fixed function mode, seriously limiting its adaptability to different types of problems. In this paper, a novel establishing method for 1-dimensional (1-D) elementary dependent function is proposed, called standard interval transformation. The remarkable advantage of this method lies in that elementary dependent function can be established based on different basis functions, leading to different function modes without influencing its inherent properties. Due to the high flexibility, the proposed method can significantly improve the adaptability of elementary dependent function to different types of real-world applications, thus supplying more reasonable references for the decision-making.

Keywords Decision making · Superiority evaluation · Extenics · Elementary dependent function · Standard interval transformation

1 Introduction

Decision-making is a common issue widely existing in the real-world. In order to obtain references for the decision-making, it is very essential to give the evaluating conclusion of each candidate scheme. Superiority evaluation, as an important branch of the Extenics [1], is a powerful evaluating method, which has been successfully applied in a wide variety of fields, such as concept design [2], problem

L. Tang (✉) · C. Yang
Research Institute of Extenics and Innovation Method, Guangdong University of Technology, Guangzhou 510006, China
e-mail: tanglong@gdut.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_52

solving [3], data mining [4], and robot techniques [5, 6]. The most critical part of superiority evaluation is establishing elementary dependent function. An elementary dependent function aims to describe the degrees of candidate schemes satisfying or not satisfying the requirement under a given evaluation characteristic. Theoretically, elementary dependent function is usually established on the basis of several nested intervals, which are obtained according to the domain knowledge that the decision-making concerns.

For the original elementary dependent function, only 1-D finite dual-nested interval case was considered [1, 7, 8]. In recent years, numerous research achievements about elementary dependent function have emerged. For example, Li and Liu [9, 10] proposed several methods to establish elementary dependent function based on infinite intervals; Smarandache [11, 12] and his partners extended 1-D elementary dependent function to 2D and 3D; Yang and Cai [13–15] refined the division of discourse universe, and proposed triple-nested interval based elementary dependent function; Li and Liu [16] introduced uncertain theory to the elementary dependent function, forming an interval general elementary dependent function.

In above research, elementary dependent function is established based on the extension distance and place value, whose function mode is fixed under the given nested intervals, seriously limiting its adaptability to different types of problems. In fact, elementary dependent function should be determined by not only the nested intervals, but also the behavior mode of the evaluation which may be linear or variously nonlinear. It is desirable that, the function mode of elementary dependent function is changeable to adapt to various types of problems. Therefore, the key of elementary dependent function lies in supplying an establishing method rather than a definite math form. In this paper, standard interval transformation method is proposed (in this paper, only 1-D elementary dependent function is considered). The method establishes elementary dependent function using interval mapping and linear transformations, by introducing basis function, standard function, standard points and standard intervals. Because the established elementary dependent function is a derivative of the basis function which can be artificially changed, the function mode becomes very flexible.

The rest of this paper is organized as follows. Section 2 introduces basic theory involving basis function, standard point, standard interval and standard interval transformation. Complete procedures of standard interval transformation method are briefly stated in Sect. 3. Section 4 investigates performance of the proposed method. Finally, conclusions are given in Sect. 5.

2 Basic Theory

Generally, possible measures of a given evaluation characteristic can be described by a real number field. In order to establish 1-D elementary dependent function, the real number field is divided into several different sub-fields, shown as Fig. 1. The positive

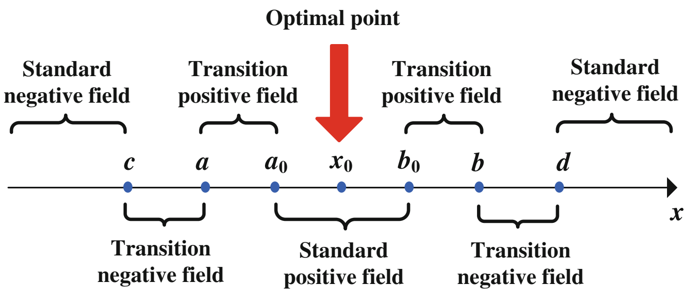


Fig. 1 Illustration of the sub-fields

field $\langle a, b\rangle$ denotes the measures satisfying the requirement, which can be further divided into the standard positive field $\langle a_0, b_0\rangle$ (the optimal point has the highest degree satisfying the requirement) and the transition positive field $\langle a, a_0\rangle \cup \langle b_0, b\rangle$, respectively representing different degrees satisfying the requirement. Similarly, the negative field $\langle -\infty, a\rangle \cup \langle b, +\infty\rangle$ denotes the measures not satisfying the requirement, which can also be further divided into the standard negative field $\langle -\infty, c\rangle \cup \langle d, +\infty\rangle$ and the transition negative field $\langle c, a\rangle \cup \langle b, d\rangle$, respectively representing different degrees not satisfying the requirement. The elementary dependent function is established in the basis of three nested intervals:

$$k(x) = k\left(x_0, X_0, X, \hat{X}\right)(x) \tag{1}$$

$$\begin{cases} X_0 = \langle a_0, b_0\rangle \\ X = \langle a, b\rangle \\ \hat{X} = X \cup X_- = \langle c, d\rangle \end{cases} \tag{2}$$

where x_0 is optimal point; X_0 is standard positive field; X is positive field; $\hat{X}$ is the union of positive field and transition negative field. Elementary dependent function $k(x)$ satisfies several following properties:

(1) $k(x)$ successively increases in $(-\infty, x_0]$;
(2) $k(x)$ successively decreases in $[x_0, +\infty)$;

(3) $$x \in X_0, 1 \le k(x) \le k(x_0);$$

(4) $$x \in X_+ = X - X_0, 0 \le k(x) \le 1;$$

(5) $$x \in X_- = \hat{X} - X, -1 \le k(x) \le 0;$$

(6) $$x \in \bar{X}_0 = R - \hat{X}, k(x) \le -1;$$

(7) When $x = a_0 \vee b_0$, $k(x) = 1$;
(8) When $x = a \vee b$, $k(x) = 0$;
(9) When $x = c \vee d$, $k(x) = -1$;

where X_+ denotes transition positive field; X_- denotes transition negative field; $\bar{X}_0$ denotes standard negative field.

2.1 Basis Function and Standard Point

Give a continuous function $f(x)$ defined in $[0, +\infty)$, satisfying following three conditions:

(1) $f(x)$ successively decreases in $[0, +\infty)$;
(2) $f(0) > 1$;
(3) $\exists\, p > 0$, when $x > p, f(x) < -1$;

Figure 2a shows the image of $f(x)$. The function values of points u, v and w are respectively 1, 0 and −1. Now define another function $F(x) = f(|x|)$. Obviously, $F(x)$ can be illustrated by Fig. 2b.

If a problem satisfies:

(1) Optimal point is $x_0 = 0$;
(2) Standard positive field is $X_0 = \langle -u, u \rangle$;
(3) Positive field is $X = \langle -v, v \rangle$;
(4) Transition negative field is $X_- = \langle -w, -v \rangle \cup \langle v, w \rangle$;

then according to the required properties of 1-D elementary dependent function, $F(x)$ can be used as the elementary dependent function $k(x)$. Here, function $f(x)$ is basis function; $F(x)$ is standard function; and points $-u$, $-v$, $-w$, 0, u, v, w are called standard points of $F(x)$.

This example, however, is just a special case. It is necessary to develop a reasonable approach that can extend above conclusion to more general cases.

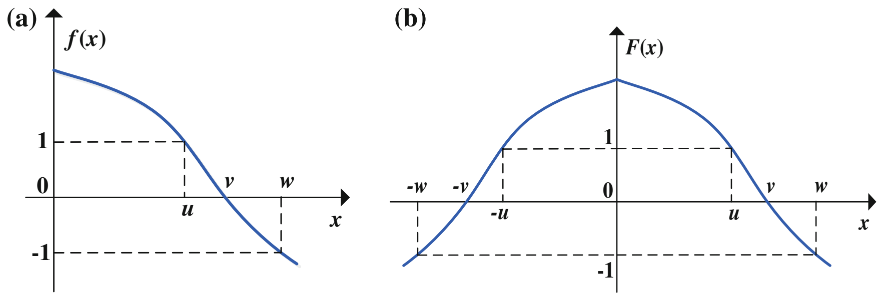


Fig. 2 Illustration of $f(x)$ and $F(x)$. **a** Illustration of $f(x)$. **b** Illustration of $F(x)$

2.2 Standard Interval Transformation

Obviously, the standard points divide the x-axis into eight different parts (Fig. 2b), called standard intervals. Similarly, endpoints a_0, b_0, a, b, c, d and optimal point x_0 also divide the x-axis into at most eight (if there is no common endpoint) different parts, called real intervals. Our purpose is to utilize the standard function $F(x)$ to establish elementary dependent function $k(x)$. In order to achieve this, a mapping relationship between real intervals and standard intervals needs to be established, shown as Fig. 3. In fact, each real interval denotes a special significance, so it should be mapped to the standard interval with equivalent significance. For instance, real interval $\langle -\infty, -c\rangle$ and standard interval $\langle -\infty, -w\rangle$ both stand for the left part of standard negative field, so $\langle -\infty, -c\rangle$ should be mapped to $\langle -\infty, -w\rangle$. Figure 4 shows the two different cases of such mapping; the first picture is the case that there is no common endpoint; the second picture is the case that $a_0 = a$ and $b = d$.

According to the mapping relationships, each real interval can be converted to corresponding standard interval by a linear transformation. A general expression of such linear transformation is:

$$T_i\left(I_i^R\right) = K_i I_i^R + Q_i \rightarrow I_i^F \tag{3}$$

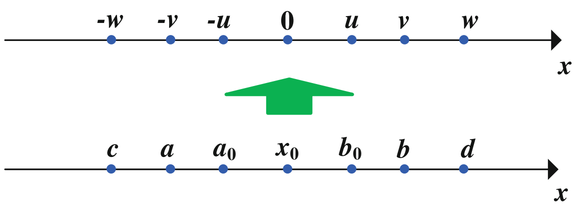


Fig. 3 Standard interval and real interval

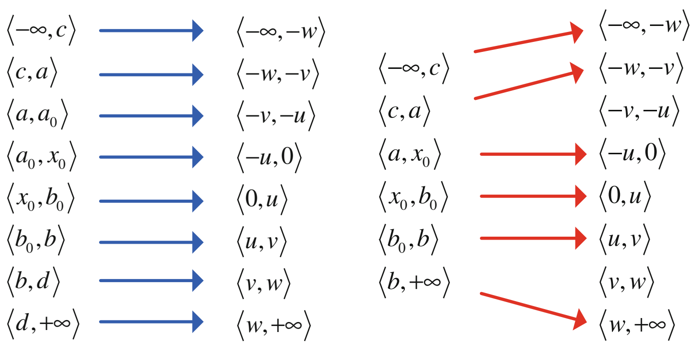


Fig. 4 Cases of mapping relationships between standard intervals and real intervals

where I_i^R denotes the ith real interval; I_i^F is the standard interval corresponding to I_i^R; $K_i > 0$ is used to increase or reduce the length of I_i^R; Q_i is a constant and used to move I_i^R along the x-axis; Values of K_i and Q_i can be determined using the end-points of real interval. Based on linear transformations, the basis function can be used for establish the elementary dependent function:

$$k(x) = F(K_i x + Q_i) = f(|K_i x + Q_i|), x \in I_i^R \tag{4}$$

In order to definitely explain standard interval transformation method, the parts on the left side of optimal point x_0 are illustrated. The right parts can be operated similarly.

(1) Considering the left part of standard negative field, denoted by $\bar{X}_0^l = \langle -\infty, c\rangle$, $I_1^R = \langle -\infty, c\rangle$ and $I_1^F = \langle -\infty, -w\rangle$ are both indefinite, so $K_1 = 1$; according to endpoint c, we have,

$$c + Q_1 = -w \Rightarrow Q_1 = -w - c \tag{5}$$

$$k(x) = F(x + Q_1) = f(|x + Q_1|), x \in \bar{X}_0^l \tag{6}$$

(2) For the left part of transition negative field $X_-^l = \langle c, a\rangle$, there are two cases:

(a) $c < a$: $I_2^R = \langle c, a\rangle$ and $I_2^F = \langle -w, -v\rangle$; according to endpoints c and a, we have,

$$\begin{pmatrix} c & 1 \\ a & 1 \end{pmatrix}\begin{pmatrix} K_2 \\ Q_2 \end{pmatrix} = \begin{pmatrix} -w \\ -v \end{pmatrix} \Rightarrow \begin{pmatrix} K_2 \\ Q_2 \end{pmatrix} = \begin{pmatrix} c & 1 \\ a & 1 \end{pmatrix}^{-1}\begin{pmatrix} -w \\ -v \end{pmatrix} \tag{7}$$

$$k(x) = F(K_2 x + Q_2) = f(|K_2 x + Q_2|), x \in X_-^l \tag{8}$$

(b) $c = a$:

(i) $X_-^l = [c, c] = \{c\}$, $k(c) = -1$;
(ii) $X_-^l = (c, c] = \phi$, neglect;
(iii) $X_-^l = [c, c) = \phi$, neglect;
(iv) $X_-^l = (c, c) = \phi$, neglect;

(3) Likewise, for the left part of transition positive field $X_+^l = \langle a, a_0\rangle$, there are also two cases:

(a) $a < a_0$: $I_3^R = \langle a, a_0\rangle$ and $I_3^F = \langle -v, -u\rangle$; according to endpoints a and a_0, we have,

$$\begin{pmatrix} a & 1 \\ a_0 & 1 \end{pmatrix}\begin{pmatrix} K_3 \\ Q_3 \end{pmatrix} = \begin{pmatrix} -v \\ -u \end{pmatrix} \Rightarrow \begin{pmatrix} K_3 \\ Q_3 \end{pmatrix} = \begin{pmatrix} a & 1 \\ a_0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} -v \\ -u \end{pmatrix} \tag{9}$$

$$k(x) = F(K_3 x + Q_3) = f(|K_3 x + Q_3|), x \in X_+^l \quad (10)$$

(b) $a = a_0$:

(i) $X_+^l = [a, a] = \{a\}$, $k(a) = 0$;
(ii) $X_+^l = (a, a] = \phi$, neglect;
(iii) $X_+^l = [a, a) = \phi$, neglect;
(iv) $X_+^l = (a, a) = \phi$, neglect;

(4) Considering the left part of standard positive field $X_0^l = \langle a_0, x_0 \rangle$, there are still two cases:

(a) $a_0 < x_0$: $I_4^R = \langle a_0, x_0 \rangle$ and $I_4^F = \langle -u, 0 \rangle$; according to endpoint a_0 and optimal point x_0, we have,

$$\begin{pmatrix} a_0 & 1 \\ x_0 & 1 \end{pmatrix} \begin{pmatrix} K_4 \\ Q_4 \end{pmatrix} = \begin{pmatrix} -u \\ 0 \end{pmatrix}, \Rightarrow \begin{pmatrix} K_4 \\ Q_4 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ x_0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -u \\ 0 \end{pmatrix} \quad (11)$$

$$k(x) = F(K_4 x + Q_4) = f(|K_4 x + Q_4|), x \in X_0^l \quad (12)$$

(b) $a_0 = x_0$: $X_0^l = \langle a_0, a_0 \rangle = \phi$, neglect;

Finally, for the optimal point x_0, $k(x_0) = F(0) = f(0)$.

3 Establishing Elementary Dependent Function by Standard Interval Transformation

This paper aims to use standard interval transformation approach to establish 1-D elementary dependent function. The main procedures are introduced as follows:

(1) Given a problem, determine the optimal point x_0, standard positive field X_0, positive field X and transition negative field X_-, shown as Eq. (2);
(2) According to priori knowledge of the given problem and related domain knowledge, establish an elementary function (satisfying the three conditions stated in Sect. 2.1) as the basis function $f(x)$; Table 1 lists several basis functions with common modes;
(3) Establish standard function $F(x) = f(|x|)$, and find out the standard points ($-u$, $-v$, $-w$, 0, u, v, w) of $F(x)$;
(4) Obtain the standard intervals and real intervals, and establish the mapping relationship between them;
(5) Use standard interval transformation method (introduced in Sect. 2.2) to establish the elementary dependent function of the different non-overlapped parts $\bar{X}_0^l$, X_-^l, X_+^l, X_0^l, X_0^r, X_+^r, X_-^r, $\bar{X}_0^r$ respectively;

Table 1 Basis functions with common modes

Basis function	Expression
Linear mode	$f(x) = 2 - x$
Square mode	$f(x) = 2 - x^2$
Square root mode	$f(x) = 2 - 3\sqrt{x}$
Fractional mode	$f(x) = \frac{2(1 - 3x^2)}{1 + x^2}$
Exponential mode	$f(x) = 4e^{-x^2} - 2$
Logarithmic mode	$f(x) = 2 - \ln(10x + 1)$
Arc-tangent mode	$f(x) = 2 - 4\arctan(x^2)$

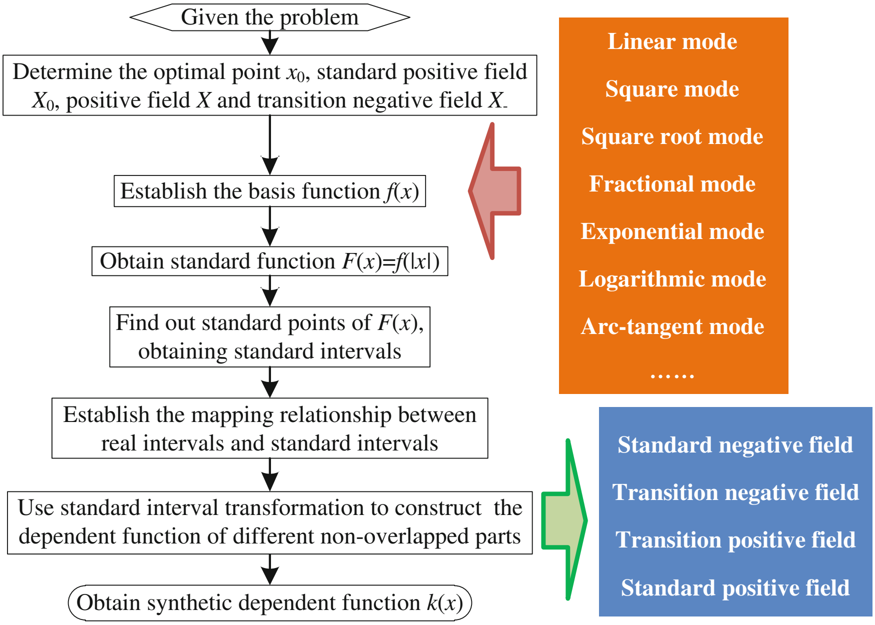


Fig. 5 Flowchart of standard interval transformation approach

(6) Obtain synthetic elementary dependent function $k(x)$.

Corresponding flowchart of above procedures is shown in Fig. 5.

4 Performance Verification

In this section, three representative cases are investigated to verify the performance of the proposed method. In each case, two different basis function modes are carried out.

Case I: X_0, X and $\hat{X}$ have no common endpoint.

$$x_0 = 1,\ X_0 = [1,3],\ X = [0,5],\ \hat{X} = (-2,7);$$

Then, we have:

$$\begin{aligned} &X_0^l = [1,1), x_0 = 1,\ X_0^r = (1,3];\\ &X_+^l = [0,1),\ X_+^r = (3,5];\\ &X_-^l = (-2,0),\ X_-^r = (5,7);\\ &\bar{X}_0^l = (-\infty, -2],\ \bar{X}_0^r = [7, +\infty); \end{aligned}$$

Two elementary dependent functions by linear and exponential mode are illustrated in Fig. 6a, b respectively. Clearly, in terms of monotonicity and function values on endpoints, two curves exhibit rigid agreement, and both satisfy the properties of elementary dependent function. Because optimal point is overlapped with the left endpoint of X_0, a discontinuity emerges here. On the other hand, varying pattern of Fig. 6a is absolutely linear, while Fig. 6b makes a difference, remarkably in standard positive field and standard negative field.

Case II: X_0 and X have a common endpoint; X and $\hat{X}$ have a common endpoint.

$$x_0 = 2.4,\ X_0 = [2.0, 2.5],\ X = (1.7, 2.5],\ \hat{X} = [1.7, 3.0);$$

Then, we have:

$$\begin{aligned} &X_0^l = [2.0, 2.4),\ x_0 = 2.4, X_0^r = (2.4, 2.5];\\ &X_+^l = (1.7, 2.0),\ X_+^r = (2.5, 2.5];\\ &X_-^l = [1.7, 1.7],\ X_-^r = (2.5, 3.0);\\ &\bar{X}_0^l = (-\infty, 1.7),\ \bar{X}_0^r = [3.0, +\infty); \end{aligned}$$

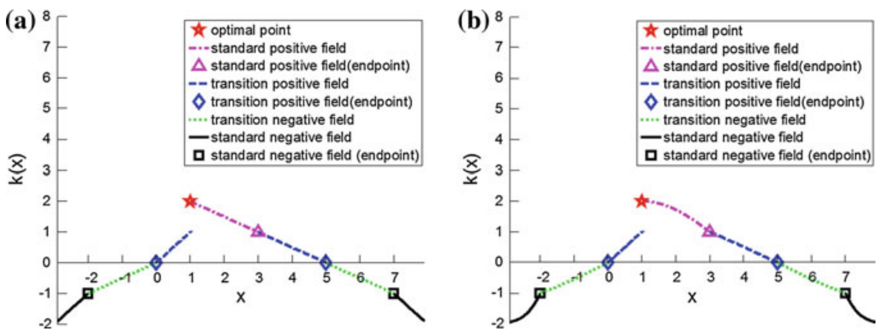


Fig. 6 Illustration of $k(x)$ in Case I by different basis function modes. **a** Linear mode. **b** Exponential mode

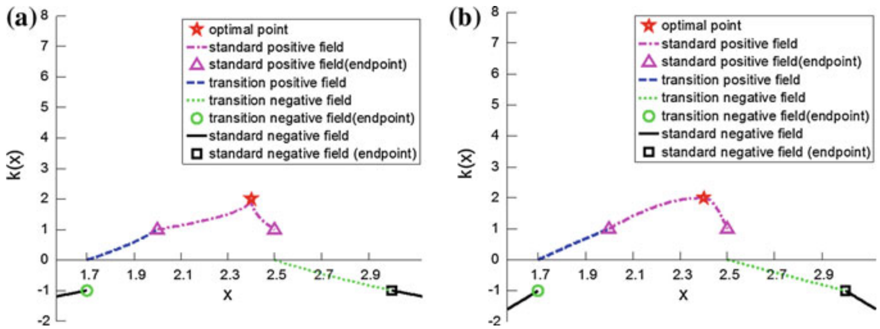


Fig. 7 Illustration of $k(x)$ in Case II by different basis function modes. **a** Square root mode. **b** Fractional mode

In this case, common endpoints are considered. Figure 7 shows the two elementary dependent functions by square root and fractional mode. At common endpoint 1.7, X is open and $\hat{X}$ is close. So endpoint 1.7 belongs to transition negative field, and its function value is -1. At common point 2.5, however, X_0 and X are close. Point 2.5 belongs to standard positive field, so its function value is 1. It also can be clearly see that, due to the diversity of basis function modes, the two curves have quite different convexities.

Case III: X_0, X and $\hat{X}$ have a common endpoint.

$$x_0 = 4,\ X_0 = (3, 7),\ X = [3, 8],\ \hat{X} = [3, 10);$$

Then, we have:

$$\begin{aligned} X_0^l &= (3, 4),\ x_0 = 4, X_0^r = (4, 7); \\ X_+^l &= [3, 3],\ X_+^r = [7, 8]; \\ X_-^l &= [3, 3),\ X_-^r = (8, 10); \\ \bar{X}_0^l &= (-\infty, 3),\ \bar{X}_0^r = [10, +\infty); \end{aligned}$$

In this case is a common endpoint involving X_0, X and $\hat{X}$. Images of two elementary dependent functions by square and logarithmic mode are given by in Fig. 8a, b. At the common endpoint 3, X_0 is open, and X and $\hat{X}$ are close. So endpoint 3 belongs to standard positive field and its function value is 0. Likewise, varying patterns of the two curves are also remarkably different from each other.

Case IV: This case considers a real-world application. Given a batch of parts, the objective is to establish the elementary dependent function of their 1-D sizes (size is the evaluation characteristic). x_0 is the ideal size; X_0 is satisfied size interval; X is acceptable size interval; X_- is transition unacceptable size interval, including the sizes that although are unacceptable but may become acceptable by reprocessing the corresponding parts. Definitely,

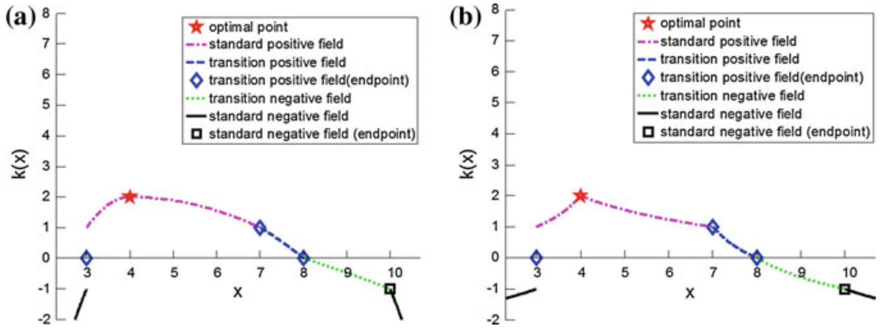


Fig. 8 Illustration of $k(x)$ in Case III by different basis function modes. **a** Square mode. **b** Logarithmic mode

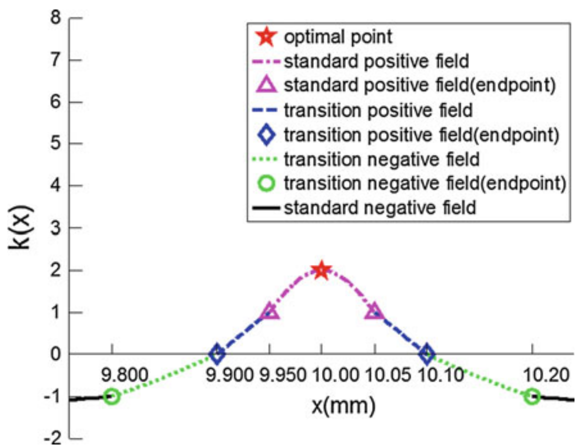
$$x_0 = 10\text{ mm}, X_0 = [9.95, 10.05]\text{ mm}, X = [9.9, 10.1]\text{ mm}, \hat{X} = [9.8, 10.2]\text{ mm}$$

Then, we have:

$$\begin{aligned} X_0^l &= [9.95, 10)\text{ mm}, x_0 = 10\text{ mm}, X_0^r = (10, 10.05]\text{ mm}; \\ X_+^l &= [9.9, 9.95)\text{ mm}, X_+^r = (10.05, 10.1]\text{ mm}; \\ X_-^l &= [9.8, 9.9)\text{ mm}, X_-^r = (10.1, 10.2]\text{ mm}; \\ \bar{X}_0^l &= (0, 9.8)\text{ mm}, \bar{X}_0^r = (10.2, +\infty)\text{ mm}; \end{aligned}$$

When the number of parts is large, the size subjects to normal distribution generally. Therefore, exponential mode is supposed to be most reasonable for this case. Obtained elementary dependent function is shown as Fig. 9, and it can well describe the degree to which the size of a part satisfies the given requirement.

Fig. 9 Illustration of $k(x)$ in Case IV



5 Conclusions

This paper proposes a standard interval transformation method to establish 1-D elementary dependent function. Differently, this method replaces extension distance and place value used conventionally to improve the establishing flexibility and adaptability. Through respectively constructing mappings from each real interval to each standard interval, elementary dependent function can be deduced from the basis function.

In real-world application, user can obtain more suitable elementary dependent function by appropriately selecting the basis function, further supplying more reasonable reference for the decision-making. In future work, the proposed standard interval transformation method will be promoted to multi-dimensional elementary dependent functions.

Acknowledgments This paper is sponsored by National Natural Science Foundation Project (61503085); National Natural Science Foundation Project (61273306); Science and Technology Planning Project of Guangdong Province (2012B061000012; "Strengthening school by innovation" Project from Department of Education of Guangdong Province (261555116).

References

1. Yang, C.Y., Cai, W.: Extension Engineering. Science Press, Beijing (2007)
2. Zhao, Y.W., Su, N.: Extension Design. Science Press, Beijing (2010)
3. Li, L.X., Yang, C.Y., Li, H.W.: Extension Strategy Generating System. Science Press, Beijing (2006)
4. Cai, W., Yang, C.Y., Chen, W.W., Li, X.S.: Extension Set and Extension Data Mining. Science Press, Beijing (2008)
5. Vladareanu, L., Sandru, O.I., Vladareanu, V.: The Robot Real Time Control using the Extenics Multidimensional Theory, Recent Advances in Robotics, Aeronautical and Mechanical Engineering, 2013, pp. 17–24
6. Vladareanu, V., Schiopu, P., Vladareanu, L.: Theory and application of extension hybrid force-position control in robotics, U.P.B. Sci. Bull. Ser. A **76**(3), 43–54 (2014)
7. Li, Q.X., Liu, S.F.: A method to construct the general location value and general elementary dependent function. Syst. Eng. **24**(6), 116–118 (2006)
8. Li, Q.X., Liu, S.F.: The Method to Construct Elementary Dependent Function Based on the Interval Distance and Side-Distance. J. Harbin Inst. Technol.: Sci. Ed. **38**(7), 1097–1100 (2006)
9. Li, Q.X.: The method to construct elementary dependent function based on infinite interval. Math. Practice Theory **39**(4), 142–146 (2009)
10. Li, Q.X.: Method to construct the elementary dependent function whose discussion fields are negative infinite interval and whole real field. Syst. Eng.-Theory Pract. **32**(12), 2740–2744 (2012)
11. Smarandache, F.: Generalizations of the distance and dependent function in extenics to 2D, 3D and n-D. Global J. Sci. Frontier Res. (GJSFR) **12**(8), 47–60 (2012)
12. Sandru, O.I., Vladareanu, L., Schiopu, P., Vladareanu, V., Sandru, A.: Multidimensional extenics theory, U.P.B. Sci. Bull. Series A **75**(1), 3–12 (2013)

13. Yang, C.Y., Cai, W.: Extenics: Theory, Method and Application. Science Press, Beijing (2013)
14. Yang, C.Y., Cai, W.: Extenics. Science Press, Beijing (2014)
15. Yang, C.Y., Cai, W.: Recent research progress in dependent functions in extension sets. J. Guangdong Univ. Technol. **29**(2), 7–14 (2012)
16. Li, Q.X., Liu, S.F.: The method to construct interval general elementary dependent function. Syst. Eng.-Theory Pract. **27**(6), 173–176 (2007)

Algorithm Research on Selection of Extension Transformation

Xiao-mei Li, Chun-yan Yang and Wei-hua Li

Abstract Using extension transformation and extension logic theory we can identify the solution transformations to solve the contradiction problems. But the number of the solution transformation set is so large that it will produce explosion problems in the number. The previously often used methods limit the number of the transformed characteristics or limit the magnitude of the transformation. But it is difficult to get a good solution transformation through these simple restrictions. The paper defines the difficulty coefficient and effect coefficient of the extension transformation. The genetic algorithm combined with the threshold of transformation difficulty coefficient can prevent explosion in the number of the extension transformation sets. And using the threshold of transformation effect coefficient, algorithm can guide the direction of extension transformation. Lastly, the example shows that the algorithm reduce the number of extension transformation greatly, and gets the solutions transformation less costly but better transformation effect. The idea provides a new basis to extension transformation selection strategy, and provides a new reference to solve the contradiction problems in the management and decision-making frontiers.

Keywords Extension transformation · Genetic algorithm · Extension transformation selection strategy

---

X. Li (✉) · W. Li
Faculty of Computer, Guangdong University of Technology,
Guangzhou, Guangdong, China
e-mail: lixm@gdut.edu.cn

X. Li · C. Yang · W. Li
Institute of Extenics and Innovation Methods, Guangdong University
of Technology, Guangzhou, Guangdong, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_53

1 Introduction

Modern society generates massive amounts of data with unimaginable speed. In addition, our living environment is constantly changing. Facing these continuously growing data and changing environment, we need to find a suitable data processing tools. The extension data mining techniques can identify changing knowledge from the data. So extension data mining [1] has been widely used in all walks of life. It is the product of the extension [2, 3] and data mining [4]. It discusses the use of the extension method and data mining technology, to mine the transformation of knowledge to solve the problem of the problem based on database.

But how to make use of these data and information, and design the appropriate extension strategy generating system [1, 5–10], to solve the contradiction problems of innovative approaches and management and decision-making frontiers. The key that Extenics solve the contradiction problems is the implementation of extension transformation. Although the results of the extension analysis modules give a general direction of on where to go to transform and which options to choose from extension transformation, but the basic transformation, transformation operator and combination still cause an explosion problems in the number of extension transformation set. Therefore, the research of selecting the extension transformation strategy is the major problems and the bottleneck problem in Extenics.

Research on extension transformation selection mechanism [11–13] is just in the primary stage of the concepts and methods, and did not conduct a comprehensive and in-depth research, and did not form a set of implementation the extension transformation choice theory as well as the corresponding software module tool. The selection algorithm of extension transformation is the key to solve the complex and contradictory issues and mining extension knowledge.

Previously facing this kind of problems, our methods include restrictions on the implementation of the number of extension transformation characteristics, such as allow only three characteristics simultaneously transformation [9]; limiting the magnitude of the transformation, such as the characteristic value "poor" can only be transformed to "middle", not allowed conversion to a higher level of "good" and "excellent" [9]; reduce the conversion category, such as the extension transformation is only limited to the replacement transformation [9]. Adding to these limitations, the number of extension transformation can indeed greatly reduced, so that the algorithm can converge to give the corresponding solution transformation. But these simple restrictions are difficult to get a good solution transformation.

Based on the above reason, this paper defines the difficulty coefficient of extension transformation to prevent explosive issues of the number of extension transformation, and defines the effect coefficient of extension transforming to guide genetic algorithm the direction of extension transformation.

2 Extension Transformation Selection Strategy

Extension strategy generating system uses the extension transformation to solve the problem of contradiction. Because the quantity of extension transformation is explosive, so need to research the selection mechanism of extension strategy. First of all, extension transform selection strategy limits the number of transformation by selecting the extension transformation characteristics. According to the selected transformation characteristics, genetic algorithm can construct individual genes encoding. By setting the three genetic operators, reduce the search space of extension transformation. Definition of the extension transformation effect coefficient and the difficulty coefficient and setting the two coefficient threshold value can determine the convergence the direction of genetic algorithm.

2.1 Select the Extension Transformation Characteristics

Based on the extension analysis results of the contradictions and problems, if the number of characteristics is large, we should select properties items using attributes reduction technology to assure the retained properties items with a larger transformation effect. Through comprehensive consideration the contribution to the transformation effect of various properties to select the properties items.

2.2 Definition of Gene Coding

According to the number of selecting the extension transformation characteristics, we can determine the gene coding length. Each extension transformation characteristic corresponds to one gene locus. The value of each gene locus is determined according to its range.

2.3 Genetic Operation

The selection operation to the group is to bring a good individual to the next generation directly, or generate new offspring individual to the next generation by the pairing. The choice of operation is based on the fitness evaluation of individual in the group. The fitness evaluation of individual is defined as the comprehensive dependent degree value.

The selected probability of the i-th individual *pi* is proportional according to its fitness, and the formula used is as follows:

$$pi = fi / \sum fi$$

In which *fi* means the fitness evaluation of the i-th individual.

The crossover operator is acted on the group to generate new individual. The crossover means that the partial structure of the two parent individuals is replaced by the structure of the recombination to form the new individual. Crossover operator is the core of genetic Algorithm.

The mutation operator means the genetic values of certain genes in the individual of the group are changed to obtained from the next generation individual according to mutation probability.

2.4 Definition of Extension Transformation Effect Coefficient

Firstly, we define the transformation effect coefficient (EC). The coefficient is defined mainly through dependent degree or comprehensive dependent degree. The effect coefficient of the single characteristic Extension transformation is defined as the difference between the dependent degree value before and after the extension transformation. The effect coefficient of the multi-characteristic Extension transform is defined as the difference between the comprehensive dependent degree value before and after the extension transformation. In the case of mining extension knowledge from the knowledge base, we can also take advantage of the magnitude of the support and confidence of the knowledge to define the transformation effect coefficient.

Set the effect coefficient threshold (EC') of single- characteristic extension transformation, so the extension transformations of the small transformation effect are discarded. The size of the threshold value can be adjusted according to the obtained solution transformation.

2.5 Definition of Extension Transformation Difficulty Coefficient

The difficulty coefficient of the single characteristic Extension transformation is defined according to the field of knowledge and expertise. And we should build a single characteristic extension transformation difficulty coefficient set.

Set extension transformation difficulty coefficient threshold (DC'), the extension transformations with the larger difficulty coefficient are discarded. The size of the threshold value can be adjusted according to the obtained solution transformations.

The definitions of the difficulty coefficient of multi-characteristic extension transformation are much more complex than single characteristic. Because there is a certain correlation between each characteristic, generally it is believed that the multi-characteristic transformation difficulty coefficient is more than the addition of the single characteristic transformation difficulty coefficient. The difficulty coefficient of multi-characteristic extension transformation should add the penalty measures according to different application environment. The specific penalty is defined according to the field of knowledge and expertise.

The single characteristic transformation difficulty coefficient with a large magnitude is not a simple linear relationship with the single characteristic transformation difficulty coefficient with a small magnitude. The determination of the difficulty coefficient also can follow the definition of multi-characteristic extension transformation difficulty coefficient model.

2.6 Determination Extension Transformation Effect Coefficient and Difficulty Coefficient Threshold

If the transformation difficulty coefficient threshold is too high and transformation effect coefficient is too low, it can not solve the explosion problem of extension transformation; If the transformation difficulty coefficient threshold is too low and effect coefficient is too high, it will result in the loss of excellent solution transformations; We should set a reasonable adjustment to these two thresholds in the algorithm design process to obtain better solution transformation.

By extension transformation difficulty coefficient and effect coefficient threshold we can guide the development direction of the extension transformation in genetic algorithm. At last we can get the solutions transformation less costly but better transformation effect, to achieve the purpose to resolve the contradictions.

2.7 Termination Conditions

It is difficult to determine the evolution of termination conditions in the light of the evolution environment complex. The maximum value of individual fitness is not clear, so it is not the base of an evolution termination decision. In this paper, according to the difficulty effect of extension transformation and extension transformation effect over the genetic algorithm we choose the optimal solution.

3 Genetic Algorithm

The specific algorithm is described as follows:

At first, based on the extension analysis results we can determine the implemented transformation characteristics. If the number of characteristics that can be

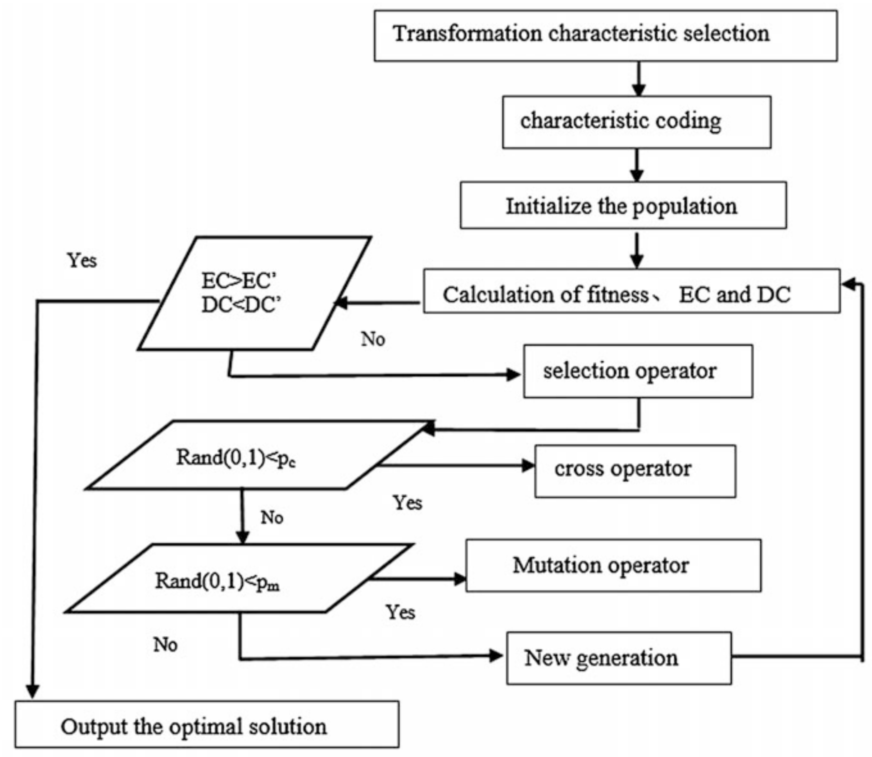


Fig. 1 Algorithm chart

implemented transformation is too large, we should use the characteristic reduction techniques to reduce the number. According to the selected characteristic we can construct single characteristic and a small magnitude transformation set. And calculate each transformation difficulty coefficient and effect coefficient.

Next, set the initial parameter in genetic algorithm. Gene coding length, the population size, cross probability pc and mutation probability pm. Set the initial transformation difficulty coefficient and effect coefficient threshold.

According to the parameter we can determine the characteristic coding and construct the initialize population. Then we can calculation the fitness, EC and DC of Individual in the population.

If the difficulty coefficient calculated in the first step is less than the corresponding threshold value, and transformation effect coefficient is greater than the corresponding threshold value, then the transformation can be into the solution transformation set. If the condition is not satisfied, then we can implement the selection operator, cross operator under the cross probability pc, mutation operator under the mutation probability pm under obtain the next generation population.

The process circle continues. If in two consecutive circles, the solutions transformation set did not continue to increase, the circle is end.

Check solutions transformation set whether there is a satisfactory result. If it is, the whole genetic algorithm is end, otherwise continue to adjust the above parameter: gene coding length, the population size, cross probability pc and mutation probability pm, transformation difficulty coefficient and effect coefficient threshold value, complement genetic algorithm again (Fig. 1).

4 Instance

In this instance through the design of the questionnaire on specialized course test, we obtain knowledge base by constructing a decision tree on specialized course test based on database. Then according to students specific contradiction problems, using extension transformation selection strategy it can get solution transformation.

4.1 Data Acquisition

We carefully designed the questionnaire on specialized course test, and received a total of 5007 results. We selected eight characteristics to analyze, including personal effort, personal interest, learning methods, specialized basic, learning atmosphere, social pressure, the exam room state and pass situation.

The information element set can be expressed as:

$$\{I_i, i = 1, 2, \ldots, 5007\}$$

In which

$$I_i = \begin{bmatrix} Student\ O_i, & hard\ c_1, & v_{i1} \\ & interest\ c_2, & v_{i2} \\ & method\ c_3, & v_{i3} \\ & base\ c_4, & v_{i4} \\ & atmosphere\ c_5, & v_{i5} \\ & pressure\ c_6, & v_{i6} \\ & state\ c_7, & v_{i7} \\ & pass\ c_8, & v_{i8} \end{bmatrix},$$

The possible values of each characteristic are:

$v_{i1}(c_1) \in \{no, middle, hard\}$ $v_{i2}(c_2) \in \{no, middle, interest\}$
$v_{i3}(c_3) \in \{no, middle, effect\}$ $v_{i4}(c_4) \in \{poor, middle, good, excellent\}$
$v_{i5}(c_5) \in \{no, middle, strong\}$ $v_{i6}(c_6) \in \{small, middle, large\}$
$v_{i7}(c_7) \in \{poor, middle, good\}$ $v_{i8}(c_8) \in \{pass, no\}$

4.2 Construct Knowledge Base

We obtain knowledge base by constructing a decision tree on database of specialized course test. The first 3 rules knowledge and corresponding support and confidence as follows:

Rule1: $[(O_i, c_4, poor) \wedge (O_i, c_2, no) \wedge (O_i, c_7, poor)] \Rightarrow (O_i, c_8, no)$
$(\mathrm{C} = 133/163, \mathrm{S} = 163/388)$
Rule2: $[(O_i, c_4, poor) \wedge (O_i, c_2, no) \wedge (O_i, c_7, middle) \wedge (O_i, c_5, no) \wedge$
$(O_i, c_6, small) \wedge (O_i, c_1, no)] \Rightarrow (O_i, c_8, no)$ $(\mathrm{C} = 3/4, \mathrm{S} = 4/6)$
Rule3: $[(O_i, c_4, poor) \wedge (O_i, c_2, no) \wedge (O_i, c_7, middle) \wedge (O_i, c_5, no) \wedge$
$(O_i, c_6, small) \wedge (O_i, c_1, hard)] \Rightarrow (O_i, c_8, pass)$ $(\mathrm{C} = 2/2, \mathrm{S} = 2/6)$

4.3 Constructed Extension Transformation Set

The implemented extension transformation includes the condition transformation and the conclusion transformation. The condition transformations include:

$T_{11}(O_i, c_1, no) = (O_i, c_1, middle)$ $T_{12}(O_i, c_1, no) = (O_i, c_1, hard)$
$T_{13}(O_i, c_1, middle) = (O_i, c_1, no)$ $T_{14}(O_i, c_1, middle) = (O_i, c_1, hard)$
$T_{15}(O_i, c_1, hard) = (O_i, c_1, no)$ $T_{16}(O_i, c_1, hard) = (O_i, c_1, middle)$
...
$T_{76}(O_i, c_7, good) = (O_i, c_7, middle)$

The conclusion transformations include:

$T_{81}(O_i, c_8, pass) = (O_i, c_8, no)$ $T_{82}(O_i, c_8, no) = (O_i, c_8, pass)$

4.4 Obtain Solutions Transformation

According to student specific contradiction problems, using extension transformation selection strategy we can get solution transformation.

For example, a specific information element can be expressed as:

$$I_i = \begin{bmatrix} Student\ O_i, & hard\ c_1, & middle \\ & interest\ c_2, & middle \\ & method\ c_3, & middle \\ & base\ c_4, & middle \\ & atmosphere\ c_5, & no \\ & pressure\ c_6, & middle \\ & state\ c_7, & middle \\ & pass\ c_8, & no \end{bmatrix},$$

Obviously this is a contradiction problem. Extension transformation in terms of the rules knowledge can be a single characteristic extension transformation, can also be a transformation combination and operation.

The difficulty coefficient of implementing a single characteristic extension transformation is easy, and the difficulty coefficient of implementation a multi-characteristic extension transformation is hard.

For a single characteristic transformation, if the extension transformation increases one level, the transformation coefficient is defined as the value of 1. If the extension transformation increases two levels, the transform coefficients is defined as the value of 3, and so on.

For a multi-characteristic transformation, we add penalty the value of 1. Its transformation difficult coefficient is defined as:

$$\begin{aligned} &\text{Multi-characteristic transformation difficult coefficient} \\ &= \sum (\text{Single characteristic transformation difficult coefficient} + 1) \end{aligned}$$

By definition transformation difficulty coefficient and setting transformation difficulty coefficient threshold combined with genetic algorithm we can avoid the explosion of the number of extension transformation.

In this example, the transformation difficulty coefficient threshold value is set as 2. According to the number of selecting the extension transformation characteristics, we can determine the gene coding length is seven. According to contradiction problem, we can construct the initial transformation individual using randomized method. In the method, randomly select a transformation characteristics, the feature value increases a grade with 0.95 probability and decreases a grade with 0.05 probability. The initial population size is 20, cross probability pc is 0.3, mutation probability pm is 0.1, and the maximum value of the iteration is 200.

Lastly, we get 40 solutions transformations. One of the solution transformations is:

$$I_i \wedge [T_{48}(O_i, c_4, middle) = (O_i, c_4, good)] \Rightarrow [T_{82}(O_i, c_8, no) = (O_i, c_8, pass)]$$

In which specialized base is transformed from "middle" to "good", the difficulty coefficient is 1, and the confident degree is 30/44 and support degree is 44/68. It is easy to find that this is a very excellent solution transformation.

If the transformation difficulty coefficient threshold value is set as 6, other parameter settings remain unchanged, we can get 402 solution transformations. One of the solution transformations is:

$$I_i \wedge [T_{46}(O_i, c_4, middle) = (O_i, c_4, excellent), T_{52}(O_i, c_5, no) = (O_i, c_5, strong)] \\ \Rightarrow [T_{82}(O_i, c_8, no) = (O_i, c_8, pass)]$$

In which "specialized base" is transformed from "middle" to "excellent", and "learning atmosphere" is transformed from "no" to "strong", the value of the difficulty coefficient is 6, and the confident degree is 23/26 and support degree is 26/153.

Obviously the solution transformation has a higher confidence, but it has too much transformation difficulty coefficient. It is difficult to be adopted as a better solution transformation.

5 Conclusion

In the paper through selection of transformation characteristics, we can control the appropriate extension transformation number of characteristics. By defining the extension transformation effect coefficient and the difficulty coefficient we can control the direction of extension transformation in genetic algorithm. The algorithm reduce the number of extension transformation greatly, and get the solutions transformation less costly but better transformation effect, to achieve the purpose to resolve the contradiction problems.

The algorithm provides a new basis to extension transformation set selection strategy, and provides a new reference to solve the contradiction problems in the management and decision-making frontiers.

Acknowledgments We thank the financial support for this research by open fund of Guangdong province key laboratory(Grant No.2011A060901001-09D) and National Science Foundation of china(Grant No.61273306 and No.61503085).

References

1. Cai, W., Yang, C., Chen, W., Li, X.: Extension Set and Extension Data Mining. Science Press, Beijing (2008)
2. Chen, W., Huang, J., Zhao, X.: Data Mining Techniques. Beijing Industry University Press, Beijing (2002)
3. Chen, W., Huang, J.: Data Warehouse and Data Mining. People's Posts and Telecommunications Press (2004)
4. Jiawei, H., Micheline, K.: Data Mining: Concepts and Techniques. Machinery Industry Press, Beijing (2007)
5. Yang, C., Cai, W.: Extension Engineering. Science Press, Beijing (2007)
6. Cai, W., Yang, C., Lin, W.: Extension Engineering Methods (2003)
7. Cai, W., Yang, C., He, B.: Extension Logic Preliminary. Science Press, Beijing (2003)
8. Chen, W., Huang, J.: From data mining to extension data mining. In: The Eleventh Chinese the Artificial Intelligence Academic Conference (2006)
9. Yang, C., Li, X., Chen, W.: Extension Data Mining Methods and their Computer Implement. Guangdong Higher Education Press, Guangzhou (2010)
10. Yang, C., Wang, G., Li, Y., Cai, W.: Study on knowledge reasoning based on extended formulas. In: International Conference on AIAI, vol. 9, pp. 797–805. Springer (2005)
11. Cai, W.: Study on Changing Knowledge and Extension Data Mining. Advances in Artificial Intelligence in China (2007), vol. 12, pp. 951–955. Beijing University of Posts and Telecommunications Press, Beijing (2007)
12. Yang, C.: Study on Knowledge Formation Based on Extension Theory. Advances in Artificial Intelligence in China (2007), vol. 12, pp. 956–961. Beijing University of Posts and Telecommunications Press, Beijing (2007)
13. Li, L., Li, H., Yang, C.: Study on the application of extenics in data mining. China Eng. Sci. **6** (7), 53–59 (2004)

Study of the Solvability of the Fuzzy Error Matrix Set Equation in Connotative Form of Type II 4

Kai-zhong Guo, Ran Li and Jian-xin Li

Abstract The concept of error matrix is presented in this paper, and the types of the fuzzy error matrix equation are presented too. The paper especially researches about the error matrix that consists of general set relations, and the solvability and solutions to it. And theorems about the necessary condition and the necessary and sufficient condition for the solvability of the error matrix equation $XA' = B$ are obtained in the paper. An example of solving this equation would be given in the last part of the paper.

Keywords Fuzzy · Error matrix · Set relation · Set equation

1 Introduction

Errors always happen in our life, work and study, sometimes with huge destructive effect, so it is necessary to study how to avoid and eliminate errors. And in order to avoid and eliminate errors, we have to study causes and laws of errors. And this paper discusses how to utilize the error matrix equation to eliminate errors.

K. Guo · R. Li (✉) · J. Li
Guangzhou Vocational College of Science and Technology, 510550 Guangzhou, China
e-mail: ranli2006@126.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_54

2 The Concept of Error Matrix

Definition 1 Let

$$\begin{pmatrix} ((U_{11} & U_{12} & \cdots & U_{1k}),\ x_{11}) & \cdots & ((U_{11} & U_{12} & \cdots & U_{1k}),\ x_{1n}) \\ ((U_{21} & U_{22} & \cdots & U_{2k}),\ x_{21}) & \cdots & ((U_{21} & U_{22} & \cdots & U_{2k}),\ x_{2n}) \\ & \cdots & & & \cdots & & & \cdots & \\ ((U_{m1} & U_{m2} & \cdots & U_{mk}),\ x_{m1}) & \cdots & ((U_{m1} & U_{m2} & \cdots & U_{mk}),\ x_{mn}) \end{pmatrix}$$

be an m × n error matrix of K elements.

Definition 2 Let

$$\begin{pmatrix} U_{20} & S_{20} & \vec{P}_{20} & T_{20} & L_{20} & y_{20} & G_{u20} \\ U_{21} & S_{21} & \vec{P}_{21} & T_{21} & L_{21} & y_{21} & G_{u21} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{2t} & S_{2t} & \vec{P}_{2t} & T_{2t} & L_{2t} & y_{2t} & G_{u2t} \end{pmatrix}$$

be a (t + 1) × 7 or m × 7 error matrix. Each element of this kind of error matrix is called a set.

Definition 3 The set relationship containing unknown sets is called set equations.

Definition 4 Let X, A′ and B be m × 7 error matrixes, so XA′ ⊇ (or other relational operators) B is named as set (matrix) equations.

3 Error Matrix Equation

Error matrix equation:

$$\text{Type I } AX \supseteq B \quad A \cdot X \supseteq B \quad A\blacktriangle X \supseteq B \quad A \vee X \supseteq B \quad A \wedge X \supseteq B$$
$$\text{Type II } XA \supseteq B \quad X \cdot A \supseteq B \quad X\blacktriangle A \supseteq B \quad X \vee A \supseteq B \quad X \wedge A \supseteq B$$

4 The Solution of Error Matrix Equation

The solution of Type II 1 XA = B

$$X = (u, v) = (U_1, S_1, \vec{P}_1, T_1, L_1, x) = \begin{pmatrix} U_{10x} & S_{10x} & \vec{P}_{10x} & T_{10x} & L_{10x} & x_{10x} & G_{10x} \\ U_{11x} & S_{11x} & \vec{P}_{11x} & T_{11x} & L_{11x} & x_{11x} & G_{11x} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{1tx} & S_{1tx} & \overrightarrow{P_{1tx}} & T_{1tx} & L_{1tx} & x_{1tx} & G_{1tx} \end{pmatrix}$$

$$A = (U_2, S_2, \vec{P}_2, T_2, L_2, x_2) = \begin{pmatrix} U_{20} & S_{20} & \vec{P}_{20} & T_{20} & L_{20} & y_{20} & G_{u20} \\ U_{21} & S_{21} & \vec{P}_{21} & T_{21} & L_{21} & y_{21} & G_{u21} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{2t} & S_{2t} & \vec{P}_{2t} & T_{2t} & L_{2t} & y_{2t} & G_{u2t} \end{pmatrix}$$

$$\begin{pmatrix} V_{201} & S_{v201} & \vec{P}_{v201} & T_{v201} & L_{v201} & y_{v201} & G_{v201} \\ U_{21j} & S_{v21j} & \vec{P}_{v21j} & T_{v21j} & L_{v21j} & y_{v21j} & G_{v21j} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{2m2m1} & S_{v2m2m1} & \vec{P}_{v2m2m1} & T_{v2m2m1} & L_{v2m2m1} & y_{v2m2m1} & G_{v2m2m1} \end{pmatrix}$$

$$B = \begin{pmatrix} (b_{11}, y_{11}) & (b_{12}, y_{12}) & \cdots & (b_{1m}, y_{1m}) \\ (b_{21}, y_{21}) & (b_{21}, y_{21}) & \cdots & (b_{2m}, y_{2m}) \\ \cdots & \cdots & \cdots & \cdots \\ (b_{m1}, y_{m1}) & (b_{m1}, y_{m1}) & \cdots & (b_{mm}, y_{mm}) \end{pmatrix}$$

Definition 5 Let

$$XA^{'} \supseteq = \begin{pmatrix} (w_{11}, z_{11}) & (w_{12}, z_{12}) & \cdots & (w_{1m}, z_{1m}) \\ (w_{21}, z_{21}) & (w_{22}, z_{22}) & \cdots & (w_{2m}, z_{2m}) \\ \cdots & \cdots & \cdots & \cdots \\ (w_{m1}, z_{m1}) & (w_{m2}, z_{m2}) & \cdots & (w_{mm}, z_{mm}) \end{pmatrix}$$

$$= \begin{pmatrix} V_{201} & S_{v201} & \overrightarrow{P}_{v201} & T_{v201} & L_{v201} & y_{v201} & G_{v201} \\ V_{21j} & S_{v21j} & \overrightarrow{P}_{v21j} & T_{v21j} & L_{v21j} & y_{v21j} & G_{v21j} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ V_{2m2m1} & S_{v2m2m1} & \overrightarrow{P}_{v2m2m1} & T_{v2m2m1} & L_{v2m2m1} & y_{v2m2m1} & G_{v2m2m1} \end{pmatrix}$$

and,

$$(w_{ij}, z_{ij}) = \left(U_{1ix} \wedge U_{2j} \quad S_{1ix} \wedge S_{2j} \quad \overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2j} \quad T_{1ix} \wedge T_{2j} \quad L_{1ix} \wedge L_{2j} \quad x_{1ix} \wedge y_{2j} \quad G_{U1ix} \wedge G_{U2j} \right)$$

So the following equation is hold.

$$\begin{pmatrix} U_{10x}\wedge U_{20} & S_{10x}\wedge S_{20} & \overrightarrow{P}_{10x}\wedge\overrightarrow{P}_{20} & T_{10x}\wedge T_{20} & L_{10x}\wedge L_{20} & x_{10x}\wedge y_{20} & G_{U10x}\wedge G_{U20} \\ U_{11x}\wedge U_{21} & S_{11x}\wedge S_{21} & \overrightarrow{P}_{11x}\wedge\overrightarrow{P}_{21} & T_{11x}\wedge T_{21} & L_{11x}\wedge L_{21} & x_{11x}\wedge y_{21} & G_{U11x}\wedge G_{U21} \\ \cdots & \cdots & \cdots & & \cdots & \cdots & \cdots \\ U_{1tx}\wedge U_{2t} & S_{1tx}\wedge S_{2t} & \overrightarrow{P}_{1tx}\wedge\overrightarrow{P}_{2t} & T_{1tx}\wedge T_{2t} & L_{1tx}\wedge L_{2t} & x_{1tx}\wedge y_{2t} & G_{U1tx}\wedge G_{U2t} \end{pmatrix}$$
$$= \begin{pmatrix} V_{201} & S_{v201} & \overrightarrow{P}_{v201} & T_{v201} & L_{v201} & y_{v201} & G_{v101} \\ V_{21j} & S_{v21j} & \overrightarrow{P}_{v21j} & T_{v21j} & L_{v21j} & y_{v21j} & G_{v11j} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ V_{2m2m1} & S_{v2m2m1} & \overrightarrow{P}_{v2m2m1} & T_{v2m2m1} & L_{v2m2m1} & y_{v2m2m1} & G_{v1m2m1} \end{pmatrix}$$

By the definition of equal matrices:if two matrices contain each other,so corresponding elements in both matrices contain each other. So $\left(w_{ij}, z_{ij}\right)\supseteq\left(b_{ij}, y_{ij}\right)$,

$$\begin{pmatrix} U_{1ix}\wedge U_{2j} & S_{1ix}\wedge S_{2j} & \overrightarrow{P}_{1ix}\wedge\overrightarrow{P}_{2j} & T_{1ix}\wedge T_{2j} & L_{1ix}\wedge L_{2j} & x_{1ix}\wedge y_{2j} & G_{U1ix}\wedge G_{U2j} \end{pmatrix}\supseteq(b_{ij}, y_{ij})$$
$$= \begin{pmatrix} V_{2ij} & S_{v2ij} & \overrightarrow{P}_{v2ij} & T_{v2ij} & L_{v2ij} & y_{v2ij} & G_{v2ij} \end{pmatrix}$$

So the following set equations are obtained:

$$\begin{gathered} U_{10x}\wedge U_{20}\supseteq V_{v20} \\ S_{10x}\wedge S_{20}\supseteq S_{v20} \\ \overrightarrow{P}_{10x}\wedge\overrightarrow{P}_{20}\supseteq\overrightarrow{P}_{v20} \\ T_{10x}\wedge T_{20}\supseteq T_{v20} \\ L_{10x}\wedge L_{20}\supseteq L_{v20} \\ x_{10x}\wedge y_{20}\supseteq y_{v20} \\ G_{U10x}\wedge G_{U20}\supseteq G_{v20} \\ \ldots\ldots \\ U_{1ix}\wedge U_{2j}\supseteq V_{v2j} \\ S_{1ix}\wedge S_{2j}\supseteq S_{v2j} \\ \overrightarrow{P}_{1ix}\wedge\overrightarrow{P}_{2j}\supseteq\overrightarrow{P}_{v2j} \\ T_{1ix}\wedge T_{2j}\supseteq T_{v2j} \\ L_{1ix}\wedge L_{2j}\supseteq L_{v2j} \\ x_{1ix}\wedge y_{2j}\supseteq y_{v2j} \\ G_{U1ix}\wedge G_{U2j}\supseteq G_{v2j} \\ \ldots\ldots \\ U_{ttx}\wedge U_{2t}\supseteq V_{v2t} \\ S_{ttx}\wedge S_{2t}\supseteq S_{v2t} \\ \overrightarrow{P}_{ttx}\wedge\overrightarrow{P}_{2t}\supseteq\overrightarrow{P}_{v2t} \\ T_{ttx}\wedge T_{2t}\supseteq T_{v2t} \\ L_{ttx}\wedge L_{2t}\supseteq L_{v2t} \\ x_{ttx}\wedge y_{2t}\supseteq y_{v2t} \\ G_{Uttx}\wedge G_{U2t}\supseteq G_{v2t} \end{gathered}$$

About the operation symbol "∧", if both sides of the equation are sets, then it means "intersection", if both sides of the equation are numbers, then it means "minimum".

As for $(U_{1\mathrm{ix}} \wedge U_{2\mathrm{j}})$ h_1 $(S_{1\mathrm{ix}} \wedge S_{2\mathrm{j}})$ h_2 $(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2j}) \vee h_3$ $(T_{1\mathrm{ix}} \wedge T_{2\mathrm{j}})$ h_4 $(L_{1\mathrm{ix}} \wedge L_{2\mathrm{j}})$ h_5 $(x_{1\mathrm{ix}} \wedge y_{2\mathrm{j}})$ h_6 $(G_{U1\mathrm{ix}} \wedge G_{\mathrm{U2j}})$, with "$h_i$, i = 1, 2, …6", means that elements have been computed could "compose" a complete matrix element (proposition). The mode of combination depends on different situations. One way is to constitute a new set of error elements or error logic proposition by parameter. And this way is called the multiplication of m × 7 error matrix.

Since what we required in solving practical problems is not $X_iA' = B$, but $X_iA' \supseteq B$, So we find a more general model of error matrix equation, that is to find out the solution of type II 1 equation $XA' \supseteq B$.

Theorem 1 *The sufficient and necessary condition for the solvability of error matrix equation* $XA \supseteq B$ *is the solvability of* $X_iA \supseteq B_i$, *i = (1, 2, ……, m2).*

Proof Suppose $XA' \supseteq B$ has solvability, it is can be known by the definitions of $XA' \supseteq B$ and $X_iA' \supseteq B_i$, i = (1, 2, …, m2) that they are the equivalent equations, so it is necessary for $X_iA' \supseteq B_i$, i = (1, 2, …, m2) has solvability; Otherwise, if the solvability of $X_iA' \supseteq B_i$, i = (1, 2, …, m2) exists, similarly it does for $XA' \supseteq B$.

Thereout, we use the method of discussing the solvability of $X_iA' = B_i$, i = (0, 1, 2, …, m2) to discuss the solution of $XA' = B$.

Then in $X_iA' \supseteq B_i$, we can get

$$\begin{aligned}
&\left(U_{1ix} \quad S_{1ix} \quad \overrightarrow{P}_{1ix} \quad T_{1ix} \quad L_{1ix} \quad v_{1ix} \quad G_{u1ix}\right)A' \supseteq \\
&(U_{1ix} \wedge U_{20}) \vee (S_{1ix} \wedge S_{20}) \vee \left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{20}\right) \vee (T_{1ix} \wedge T_{20}) \vee (L_{1ix} \wedge L_{20}) \\
&\vee (x_{1ix} \wedge x_{20}) \vee (G_{U1ix} \wedge G_{U20}) \\
&\ldots\ldots \\
&\left(U_{1ix} \wedge U_{2j}\right) \vee \left(S_{1ix} \wedge S_{2j}\right) \vee \left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2j}\right) \vee \left(T_{1ix} \wedge T_{2j}\right) \vee \left(L_{1ix} \wedge L_{2j}\right) \\
&\vee \left(x_{1ix} \wedge x_{2j}\right) \vee \left(G_{U1ix} \wedge G_{U2j}\right) \\
&\ldots\ldots \\
&(U_{1ix} \wedge U_{2m1}) \vee (S_{1ix} \wedge S_{2m1}) \vee \left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2m1}\right) \vee (T_{1ix} \wedge T_{2m1}) \\
&\vee (L_{1ix} \wedge L_{2m1}) \vee (x_{1ix} \wedge x_{2m1}) \vee (G_{U1ix} \wedge G_{U2m1}) \\
&\quad \supseteq \left((b_{i1}, y_{i1}) \quad (b_{i2}, y_{i2}) \quad \cdots \quad (b_{im1}, y_{im1}) \right)
\end{aligned}$$

Namely,

$$\left(U_{1ix} \wedge U_{20}\right) \vee \left(S_{1ix} \wedge S_{20}\right) \vee \left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{20}\right) \vee \left(T_{1ix} \wedge T_{20}\right)$$
$$\vee \left(L_{1ix} \wedge L_{20}\right) \vee \left(x_{1ix} \wedge x_{20}\right) \vee \left(G_{U1ix} \wedge G_{U20}\right) \supseteq$$
$$\left(V_{20} \quad S_{v20} \quad \overrightarrow{P}_{v20} \quad T_{v20} \quad L_{v20} \quad y_{v20} \quad G_{v20}\right)$$
$$\cdots\cdots$$
$$\left(U_{1ix} \wedge U_{2j}\right) \vee \left(S_{1ix} \wedge S_{2j}\right) \vee \left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2j}\right) \vee \left(T_{1ix} \wedge T_{2j}\right) \vee \left(L_{1ix} \wedge L_{2j}\right)$$
$$\vee \left(x_{1ix} \wedge x_{2j}\right) \vee \left(G_{U1ix} \wedge G_{U2j}\right) \supseteq$$
$$\left(V_{2j} \quad S_{v2j} \quad \overrightarrow{P}_{v2j} \quad T_{v2j} \quad L_{v2j} \quad y_{v2j} \quad G_{v2j}\right)$$
$$\cdots\cdots$$
$$\left(U_{1tx} \wedge U_{2m1}\right) \vee \left(S_{1tx} \wedge S_{2m1}\right) \vee \left(\overrightarrow{P}_{1tx} \wedge \overrightarrow{P}_{2m1}\right) \vee \left(T_{1tx} \wedge T_{2m1}\right) \vee \left(L_{1tx} \wedge L_{2m1}\right)$$
$$\vee \left(x_{1tx} \wedge x_{2m1}\right) \vee \left(G_{U1tx} \wedge G_{U2m1}\right) \supseteq$$
$$\left(V_{2t} \quad S_{v2t} \quad \overrightarrow{P}_{v2t} \quad T_{v2t} \quad L_{v2t} \quad y_{v2t} \quad G_{v2t}\right)$$

A series of set equations be obtained:

$$\begin{gathered}
\left(U_{1ix} \wedge U_{20}\right) \supseteq V_{20} \\
\left(S_{1ix} \wedge S_{20}\right) \supseteq S_{v20} \\
\left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{20}\right) \supseteq \overrightarrow{P}_{v20} \\
\left(T_{1ix} \wedge T_{20}\right) \supseteq T_{v20} \\
\left(L_{1ix} \wedge L_{20}\right) \supseteq L_{v20} \\
\left(x_{1ix} \wedge x_{20}\right) \supseteq y_{v20} \\
\left(G_{U1ix} \wedge G_{U20}\right) \supseteq G_{v20} \\
\cdots\cdots \\
\left(U_{1ix} \wedge U_{2j}\right) \supseteq V_{2j} \\
\left(S_{1ix} \wedge S_{2j}\right) \supseteq S_{v2j} \\
\left(\overrightarrow{P}_{1ix} \wedge \overrightarrow{P}_{2j}\right) \supseteq \overrightarrow{P}_{v2j} \\
\left(T_{1ix} \wedge T_{2j}\right) \supseteq T_{v2j} \\
\left(L_{1ix} \wedge L_{2j}\right) \supseteq L_{v2j} \\
\left(x_{1ix} \wedge x_{2j}\right) \supseteq y_{v2j} \\
\left(G_{U1ix} \wedge G_{U2j}\right) \supseteq G_{v2j} \\
\cdots\cdots \\
\left(U_{1tx} \wedge U_{2m1}\right) \supseteq V_{2t} \\
\left(S_{1tx} \wedge S_{2m1}\right) \supseteq S_{v2t} \\
\left(\overrightarrow{P}_{1tx} \wedge \overrightarrow{P}_{2m1}\right) \supseteq \overrightarrow{P}_{v2t} \\
\left(T_{1tx} \wedge T_{2m1}\right) \supseteq T_{v2t} \\
\left(L_{1tx} \wedge L_{2m1}\right) \supseteq L_{v2t} \\
\left(x_{1tx} \wedge x_{2m1}\right) \supseteq y_{v2t} \\
\left(G_{U1tx} \wedge G_{U2m1}\right) \supseteq G_{v2t}
\end{gathered}$$

Theorem 2 *The sufficient and necessary condition for the solvability of* $X_iA \supseteq B_i$ *is,*

$$
\begin{gathered}
U_{20} \supseteq V_{20} \\
S_{20} \supseteq S_{v20} \\
\vec{P}_{20} \supseteq \vec{P}_{v20} \\
T_{20} \supseteq T_{v20} \\
L_{20} \supseteq L_{v20} \\
x_{20} \supseteq y_{v20} \\
G_{U20} \supseteq G_{v20} \\
\cdots\cdots \\
U_{2j} \supseteq V_{2j} \\
S_{2j} \supseteq S_{v2j} \\
\vec{P}_{2j} \supseteq \vec{P}_{v2j} \\
T_{2j} \supseteq T_{v2j} \\
L_{2j} \supseteq L_{v2j} \\
x_{2j} \supseteq y_{v2j} \\
G_{U2j} \supseteq G_{v2j} \\
\cdots\cdots \\
U_{2m1} \supseteq V_{2t} \\
S_{2m1} \supseteq S_{v2t} \\
\vec{P}_{2m1} \supseteq \vec{P}_{v2t} \\
T_{2m1} \supseteq T_{v2t} \\
L_{2m1} \supseteq L_{v2t} \\
x_{2m1} \supseteq y_{v2t} \\
G_{U2m1} \supseteq G_{v2t}
\end{gathered}
$$

Proof 1 If one of the conditions above is not satisfied, for example suppose $S_{2j}(t) \supseteq S_{v2j}(t)$ is not satisfied, so in the $(S_{1ix}(t) \wedge S_{2j}(t)) = S_{v2j}(t)$, no mater what value $S_{1ix}(t)$ is, we can not get $(S_{1ix}(t) \wedge S_{2j}(t)) = S_{v2j}(t)$。

Proof 2 Since we could only take union operation between the corresponding element of A and X_i in the $X_iA' \supseteq B_i$, that is

$$
\begin{aligned}
U_{1ix} &= U_{20} \cup U_{21} \cup \cdots U_{2j} \cup \cdots \cup U_{2t} \\
S_{1ix} &= S_{20} \cup S_{21} \cup \cdots S_{2j} \cup \cdots \cup S_{2t} \\
\vec{P}_{1ix} &= \vec{P}_{20} \cup \vec{P}_{21} \cup \cdots \vec{P}_{2j} \cup \cdots \cup \vec{P}_{2t} \\
T_{1ix} &= T_{20} \cup T_{21} \cup \cdots T_{2j} \cup \cdots \cup T_{2t} \\
L_{1ix} &= L_{20} \cup L_{21} \cup \cdots L_{2j} \cup \cdots \cup L_{2t} \\
x_{1ix} &= x_{20} \cup x_{21} \cup \cdots x_{2j} \cup \cdots \cup x_{2t} \\
G_{U1ix} &= G_{U20} \cup G_{U21} \cup \cdots G_{U2j} \cup \cdots \cup G_{U2t}
\end{aligned}
$$

Then we discuss all the solutions of $X_iA' \supseteq B_i$ and $XA' \supseteq B$.

When $X(x_1, x_2, x_n)$ are obtained, we take intersection operation between X 与 Kg, rw, xq, so we can get $X'(x'_1, x'_2, \ldots, x'_n) \in X'$。

5 Examples of Error Matrix Equation

Suppose $A' = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and

$$a_{11} = \begin{pmatrix} U_{201} & S_{201} & \overrightarrow{P}_{201} & T_{201} & L_{201} & y_{201} & G_{U201} \end{pmatrix}$$
$$a_{12} = \begin{pmatrix} U_{202} & S_{202} & \overrightarrow{P}_{202} & T_{202} & L_{202} & y_{202} & G_{U202} \end{pmatrix}$$
$$a_{21} = \begin{pmatrix} U_{211} & S_{211} & \overrightarrow{P}_{211} & T_{211} & L_{211} & y_{211} & G_{U211} \end{pmatrix}$$
$$a_{22} = \begin{pmatrix} U_{212} & S_{212} & \overrightarrow{P}_{212} & T_{212} & L_{212} & y_{212} & G_{U212} \end{pmatrix}$$

When n = 11, we get the following equations:

$$U_{201} = \{u_{201}, u_{202}, \ldots, u_{20n}\}$$
$$S_{201} = \{s_{201}, s_{202}, \ldots, s_{20n}\}$$
$$\overrightarrow{P}_{201} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20n}\}$$
$$T_{201} = \{t_{201}, t_{202}, \ldots, t_{20n}\}$$
$$L_{201} = \{l_{201}, l_{202}, \ldots, l_{20n}\}$$
$$y_{201} = \{y_{201}, y_{202}, \ldots, y_{20n}\}$$
$$G_{U201} = \{g_{201}, g_{202}, \ldots, g_{20n}\}$$

When n = 9, we get the following equations:

$$U_{202} = \{u_{201}, u_{202}, \ldots, u_{20n}\}$$
$$S_{202} = \{s_{201}, s_{202}, \ldots, s_{20n}\}$$
$$\overrightarrow{P}_{202} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20n}\}$$
$$T_{202} = \{t_{201}, t_{202}, \ldots, t_{20n}\}$$
$$L_{202} = \{l_{201}, l_{202}, \ldots, l_{20n}\}$$
$$y_{202} = \{y_{201}, y_{202}, \ldots, y_{20n}\}$$
$$G_{U202} = \{g_{201}, g_{202}, \ldots, g_{20n}\}$$

When n = 10, we get the following equations:

$$
\begin{aligned}
U_{211} &= \{u_{211}, u_{212}, \ldots, u_{21n}\} \\
S_{211} &= \{s_{211}, s_{212}, \ldots, s_{21n}\} \\
\overrightarrow{P}_{211} &= \{\overrightarrow{p}_{211}, \overrightarrow{p}_{212}, \ldots, \overrightarrow{p}_{21n}\} \\
T_{211} &= \{t_{211}, t_{212}, \ldots, t_{21n}\} \\
L_{211} &= \{l_{211}, l_{212}, \ldots, l_{21n}\} \\
y_{211} &= \{y_{211}, y_{212}, \ldots, y_{21n}\} \\
G_{U211} &= \{g_{211}, g_{212}, \ldots, g_{21n}\}
\end{aligned}
$$

When n = 15, we get the following equations:

$$
\begin{aligned}
U_{212} &= \{u_{211}, u_{212}, \ldots, u_{21n}\} \\
S_{212} &= \{s_{211}, s_{212}, \ldots, s_{21n}\} \\
\overrightarrow{P}_{212} &= \{\overrightarrow{p}_{211}, \overrightarrow{p}_{212}, \ldots, \overrightarrow{p}_{21n}\} \\
T_{212} &= \{t_{211}, t_{212}, \ldots, t_{21n}\} \\
L_{212} &= \{l_{211}, l_{212}, \ldots, l_{21n}\} \\
y_{212} &= \{y_{211}, y_{212}, \ldots, y_{21n}\} \\
G_{U212} &= \{g_{211}, g_{212}, \ldots, g_{21n}\}
\end{aligned}
$$

And suppose $X = \begin{pmatrix} x_1 & x_2 \end{pmatrix}$, where

$$
\begin{aligned}
x_1 &= \begin{pmatrix} U_{10x} & S_{10x} & \overrightarrow{P}_{10x} & T_{10x} & L_{10x} & x_{10x} & G_{U10x} \end{pmatrix} \\
x_2 &= \begin{pmatrix} U_{11x} & S_{11x} & \overrightarrow{P}_{11x} & T_{11x} & L_{11x} & x_{11x} & G_{U11x} \end{pmatrix}
\end{aligned}
$$

And suppose $B' = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, where

$$
\begin{aligned}
b_{11} &= \begin{pmatrix} V_{201} & S_{v201} & \overrightarrow{P}_{v201} & T_{v201} & L_{v201} & y_{v201} & G_{V201} \end{pmatrix} \\
b_{12} &= \begin{pmatrix} V_{202} & S_{v202} & \overrightarrow{P}_{v202} & T_{v202} & L_{v202} & y_{v202} & G_{V202} \end{pmatrix} \\
b_{21} &= \begin{pmatrix} V_{211} & S_{v211} & \overrightarrow{P}_{v211} & T_{v211} & L_{v211} & y_{v211} & G_{V211} \end{pmatrix} \\
b_{22} &= \begin{pmatrix} V_{212} & S_{v212} & \overrightarrow{P}_{v212} & T_{v212} & L_{v212} & y_{v212} & G_{V212} \end{pmatrix}
\end{aligned}
$$

And when k = 7, we get the following equations:

$$V_{201} = \{u_{201}, u_{202}, \ldots, u_{20k}\}$$
$$S_{v201} = \{s_{201}, s_{202}, \ldots, s_{20k}\}$$
$$\overrightarrow{P}_{v201} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20k}\}$$
$$T_{v201} = \{t_{201}, t_{202}, \ldots, t_{20k}\}$$
$$L_{v201} = \{l_{201}, l_{202}, \ldots, l_{20k}\}$$
$$y_{v201} = \{y_{201}, y_{202}, \ldots, y_{20k}\}$$
$$G_{v201} = \{g_{201}, g_{202}, \ldots, g_{20k}\}$$

And when k = 8, we get the following equations:

$$V_{202} = \{u_{201}, u_{202}, \ldots, u_{20k}\}$$
$$S_{v202} = \{s_{201}, s_{202}, \ldots, s_{20k}\}$$
$$\overrightarrow{P}_{v202} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20k}\}$$
$$T_{v202} = \{t_{201}, t_{202}, \ldots, t_{20k}\}$$
$$L_{v202} = \{l_{201}, l_{202}, \ldots, l_{20k}\}$$
$$y_{v202} = \{y_{201}, y_{202}, \ldots, y_{20k}\}$$
$$G_{v202} = \{g_{201}, g_{202}, \ldots, g_{20k}\}$$

And when k = 2, we get the following equations:

$$V_{211} = \{u_{211}, u_{212}, \ldots, u_{21k}\}$$
$$S_{v211} = \{s_{211}, s_{212}, \ldots, s_{21k}\}$$
$$\overrightarrow{P}_{v211} = \{\overrightarrow{p}_{211}, \overrightarrow{p}_{212}, \ldots, \overrightarrow{p}_{21k}\}$$
$$T_{v211} = \{t_{211}, t_{212}, \ldots, t_{21k}\}$$
$$L_{v211} = \{l_{211}, l_{212}, \ldots, l_{21k}\}$$
$$y_{v211} = \{y_{211}, y_{212}, \ldots, y_{21k}\}$$
$$G_{v211} = \{g_{211}, g_{212}, \ldots, g_{21k}\}$$

And when k = 3, we get the following equations:

$$V_{212} = \{u_{211}, u_{212}, \ldots, u_{21k}\}$$
$$S_{v212} = \{s_{211}, s_{212}, \ldots, s_{21k}\}$$
$$\overrightarrow{P}_{v212} = \{\overrightarrow{p}_{211}, \overrightarrow{p}_{212}, \ldots, \overrightarrow{p}_{21k}\}$$
$$T_{v212} = \{t_{211}, t_{212}, \ldots, t_{21k}\}$$
$$L_{v212} = \{l_{211}, l_{212}, \ldots, l_{21k}\}$$
$$y_{v212} = \{y_{211}, y_{212}, \ldots, y_{21k}\}$$
$$G_{v212} = \{g_{211}, g_{212}, \ldots, g_{21k}\}$$

By the Theorem 2, when n = 11, we get the solution of XA′ ⊇ B, which is

$$U_{10x} = \{u_{201}, u_{202}, \ldots, u_{20n}\}$$
$$S_{10x} = \{s_{201}, s_{202}, \ldots, s_{20n}\}$$
$$\overrightarrow{P}_{10x} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20n}\}$$
$$L_{10x} = \{l_{201}, l_{202}, \ldots, l_{20n}\}$$
$$x_{10x} = \{y_{201}, y_{202}, \ldots, y_{20n}\}$$
$$G_{U10x} = \{g_{201}, g_{202}, \ldots, g_{20n}\}$$

Similarly when n = 8 we get the following equations:

$$U_{11x} = \{u_{211}, u_{212}, \ldots, u_{20n}\}$$
$$S_{11x} = \{s_{201}, s_{202}, \ldots, s_{20n}\}$$
$$\overrightarrow{P}_{11x} = \{\overrightarrow{p}_{201}, \overrightarrow{p}_{202}, \ldots, \overrightarrow{p}_{20n}\}$$
$$T_{11x} = \{t_{201}, t_{202}, \ldots, t_{20n}\}$$
$$x_{11x} = \{y_{211}, y_{212}, \ldots, y_{21n}\}$$
$$G_{U11x} = \{g_{211}, g_{212}, \ldots, g_{21n}\}$$

6 Conclusion

We get the necessary and sufficient condition for the solvability of the error matrix equation XA′ = B, and they are also can be proved by the case studies.

Acknowledgments Thanks to the support by Guangzhou Vocational College of Science and Technology of China (No. 2015ZR01 and No. 2015JG08) and the provincial education system reform project of Department of Education of Guangdong Province, which gets the financial support form Department of Finance of Guangdong Province, with the number 16 in the third category of the second tier.

Bibliography

1. Kai-zhong, G., Shi-qiang, Z.: The Introduction to Eliminating Error. South China University of Technology Press, Guangzhou (1995)
2. Hong-bin, L., Kai-zhong, G.: One-element fuzzy error-matrix. Modeling **27**(2), 33–42 (2006)
3. Hai'ou, X., Kai-zhong, G.: Research on Decomposition conversion connectors of fuzzy logic error domain. Fuzzy Syst. Math. **20**(1), 24–29 (2006)
4. Shiyong, L., Kai-zhong, G.: Substantail change of decision-making environment-mutation of fuzzy error system. Advances A **39**(4), 29–39 (2002)
5. Cheng-yi, W.: An Introduction to Fuzzy Mathematics. Beijing University of Technology Press, Beijing (1998)
6. Kai-zhong, G., Shi-qiang, Z.: Research on laws of security risk- error logic system with critical point. Model. Measur. Control **22**(1), 1–10 (2001)

A New Measure of Similarity Between Vague Sets

Bing-jiang Zhang

Abstract Vague sets, as generalized fuzzy sets, have more powerful ability to process fuzzy information than fuzzy sets. In this paper, we propose a new similarity measure based on internal difference and external difference between vague sets. The proposed similarity measure approach can better understand the relationship between vague sets in applications. Compared to existing similarity measures, the approach is far more reasonable and effective in measuring the similarity between vague sets.

Keywords Fuzzy sets · Vague sets · Similarity measure

1 Introduction

In the classical set theory introduced by Cantor, a German mathematician, values of elements in a set are only one of 0 and 1. That is, for all element, there are only two possibilities: in or not in the set. Therefore, the theory cannot handle the data with ambiguity and uncertainty.

Zadeh proposed fuzzy theory in 1965 [1]. The most important feature of a fuzzy set is that fuzzy set A is a class of objects that satisfy a certain (or several) property. In fuzzy set theory, each object $u \in U$ is assigned a single real value, called the grade of membership, between zero and one. (Here U is a classical set of objects, called the universe of discourse.) Formally, a membership function $\mu_A \colon U \to [0, 1]$ is defined for a fuzzy set A, where $\mu_A(u)$ is the grade of membership of element u in set A, for each $u \in U$, denotes the degree of membership of u in the fuzzy set A. This membership function has the following characteristics: the single degree contains the evidences for both supporting and opposing u. It cannot only represent one of the two evidences, but also represent both at the same time.

B. Zhang (✉)
School of Applied Science, Beijing Information Science and Technology University, Beijing 100192, China
e-mail: zbj2013ch@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_55

In order to deal with this problem, Gau and Buehrer proposed the concept of vague sets in 1993 [2], by replacing the value of an element in a set with a sub-interval of $[0,1]$. Namely, a truth-membership function $t_A(u)$ and a false-membership function $f_A(u)$ are used to describe the boundaries of membership degree. These two boundaries form a sub-interval $[t_A(u), 1-f_A(u)]$ of $[0,1]$, to generalize the $\mu_A(u)$ of fuzzy sets, where $t_A(u) \le \mu_A(u) \le 1-f_A(u)$. The vague set theory improves description of the objective real world, becoming a promising tool to deal with inexact, uncertain or vague knowledge. Many researchers have applied this theory to many situations, such as fuzzy control, decision-making, knowledge discovery and fault diagnosis. And the tool has presented more challenging than that with fuzzy sets theory in applications.

In intelligent activities, it is often needed to compare between two fuzzy concepts. That is, we need to check whether two knowledge patterns are identical or approximately equal, to find out functional dependence relations between concepts in a data mining system. In order to solve this problem, many measure methods have been proposed to measure the similarity between two vague sets (values). Each of them is given from different side, having its own counterexamples.

After analyzing most existing vague sets and similarity measures of vague values, this paper develops a method for the similarity measure of vague sets.

2 Preliminaries

In this section, we review some basic definitions of vague values and vague sets.

Definition 1 Let U be a space of points (objects), with a generic element of U denoted by u. A vague set A in U is characterized by a truth-membership function $t_A(u)$ and a false-membership function $f_A(u)$. $t_A(u)$ is a lower bound on the grade of membership of u derived from the evidence for u, and $f_A(u)$ is a lower bound on the negation of u derived from the evidence against u, $t_A(u)$ and $f_A(u)$ both associate a real number in the interval $[0,1]$ with each point in U, where $t_A(u) + f_A(u) \le 1$. That is $t_A\colon U \to [0,1]$ and $f_A\colon U \to [0,1]$.

This approach bounds the grade of membership of u to a subinterval $[t_A(u), 1-f_A(u)]$ of [0, 1]. We depict these ideas in Fig. 1.

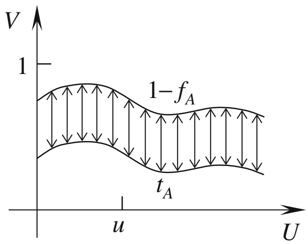


Fig. 1 A vague set

The precision of the knowledge about u is characterized by the difference $1-t_A(u)-f_A(u)$. If it is small, the knowledge about u is relatively precise. If it is large, we know correspondingly little. If $t_A(u)$ is equal to $1-f_A(u)$, the knowledge about u is exact, and the vague set theory reverts back to fuzzy set theory. If $t_A(u)$ and $1-f_A(u)$ are both equal to 1 or 0, depending on whether u does or does not belong to A, the knowledge about u is very exact and the theory reverts back to ordinary sets.

When U is continuous, a vague set A of the universe of discourse U can be represented by

$$A=\int_U [t_A(u), 1-f_A(u)]/u, u\in U, \tag{1}$$

where $t_A(u)\le \mu_A(u)\le 1-f_A(u)$.

When U is discrete, a vague set A of the universe of discourse U can be represented by

$$A=\sum_{i=1}^{n} [t_A(u_i), 1-f_A(u_i)]/u_i, \forall u_i\in U, \tag{2}$$

where $t_A(u_i)\le \mu_A(u_i)\le 1-f_A(u_i)$ and $1\le i\le n$.

Definition 2 Let x be a vague value, where $x=[t_x, 1-f_x]$. If $t_x=1$ and $f_x=0$ (i.e. $x=[1,1]$), then x is called a unit vague value.

Definition 3 Let x be a vague value, where $x=[t_x, 1-f_x]$. If $t_x=0$ and $f_x=1$ (i.e. $x=[0,0]$), then x is called a zero vague value.

Definition 4 Let x and y be two vague values, where $x=[t_x, 1-f_x]$ and $y=[t_y, 1-f_y]$. If $t_x=t_y$ and $f_x=f_y$, then the vague values x and y are called equal (i.e. $[t_x, 1-f_x]=[t_y, 1-f_y]$).

Definition 5 Let A be a vague set of $U=\{u_1, u_2 \ldots, u_n\}$. If $\forall i$, $t_A(u_i)=1$ and $f_A(u_i)=0$, then A is called a unit vague set, where $1\le i\le n$.

Definition 6 Let A be a vague set of $U=\{u_1, u_2 \ldots, u_n\}$. If $\forall i$, $t_A(u_i)=0$ and $f_A(u_i)=1$, then A is called a zero vague set, where $1\le i\le n$.

Definition 7 Let A be a vague set of $U=\{u_1, u_2 \ldots, u_n\}$. If $\forall i$, $t_A(u_i)=0$ and $f_A(u_i)=0$, then A is called an empty vague set, where $1\le i\le n$.

Definition 8 Let A and B be vague sets of universe of discourse $U=\{u_1, u_2 \ldots, u_n\}$, $A=\sum_{i=1}^{n} [t_A(u_i), 1-f_A(u_i)]/u_i$, $B=\sum_{i=1}^{n} [t_B(u_i), 1-f_B(u_i)]/u_i$. If $\forall i$, $[t_A(u_i), 1-f_A(u_i)]=[t_B(u_i), 1-f_B(u_i)]$, then the vague sets A and B are called equal.

3 Research into Similarity Measure

Currently, there have been many similarity measurements for vague sets (values).

Suppose that $x = [t_x, 1 - f_x]$ and $y = [t_y, 1 - f_y]$ are two vague values over the universe of discourse U. Let $S(x) = t_x - f_x$, $S(y) = t_y - f_y$, the M_c, M_H, M_L and M_O models are defined respectively as follows [3–6]:

$$M_c(x,y) = 1 - \frac{|S(x) - S(y)|}{2} = 1 - \frac{\left|\left(t_x - t_y\right) - \left(f_x - f_y\right)\right|}{2}, \tag{3}$$

$$M_H(x,y) = 1 - \frac{\left|t_x - t_y\right| - \left|f_x - f_y\right|}{2}, \tag{4}$$

$$\begin{aligned} M_L(x,y) &= 1 - \frac{|S(x) - S(y)|}{4} \frac{\left|t_x - t_y\right| - \left|f_x - f_y\right|}{2}, \\ &= 1 - \frac{\left|\left(t_x - t_y\right) - \left(f_x - f_y\right)\right| + \left|t_x - t_y\right| + \left|f_x - f_y\right|}{2} \end{aligned} \tag{5}$$

$$M_O(x,y) = 1 - \sqrt{\frac{\left(t_x - t_y\right)^2 + \left(f_x - f_y\right)^2}{2}}, \tag{6}$$

For the $M_C(x,y)$, $M_H(x,y)$ and $M_L(x,y)$, consider the vague values $x = [1,1]$ and $y = [a, 1-a]$, $0 \le a \le 1$. According to formulas (3)–(5), it can be checked that $M_C(x,y) = M_H(x,y) = M_L(x,y) = 0.5$. This means that x and y are equal, i.e. the $M_C(x,y)$, $M_H(x,y)$ and $M_L(x,y)$ models are rough when $x = [1,1]$ and $y = [a, 1-a]$. The $M_O(x,y)$ model does not consider whether the differences are positive or negative. On the other hand, let $x = [t_x, 1 - f_x]$ and $y = [1,1]$, the formula (6) can be written as:

$$M_O(x,y) = 1 - \sqrt{\frac{(t_x - 1)^2 + (f_x)^2}{2}} = c, \tag{7}$$

where c is the degree of similarity between the vague values x and y. According to formula (7), we can get the relational expression of f_x and t_x as follows:

$$f_x = \sqrt{2(1-c)^2 - (t_x - 1)^2}. \tag{8}$$

If c is a fixed value, then we can get a series of vague values $[t_x, 1 - \sqrt{2(1-c)^2 - (t_x - 1)^2}]$, and the degree of similarity between the vague values $[t_x, 1 - \sqrt{2(1-c)^2 - (t_x - 1)^2}]$ and $[1,1]$ is equal, i.e. the $M_O(x,y)$ model is too rough.

Hamming distance [7] and Euclidean distance [8] to measure the distances between Vague sets as follows:

1. Hamming distance is given by

$$D_H(x,y) = \frac{|t_x - t_y| + |f_x - f_y| + |(t_x - t_y) + (f_x - f_y)|}{2} \tag{9}$$

2. Euclidean distance is given by

$$D_E(x,y) = \sqrt{\frac{(t_x - t_y)^2 + (f_x - f_y)^2 + ((t_x - t_y) + (f_x - f_y))^2}{2}} \tag{10}$$

These methods also have some problems. We consider the vague values $x = [0, 1]$ and $y = [a, a], 0 \leq a \leq 1$. For the Hamming distance, it can be calculated that $D_H(x,y) = 1$. This means that the Hamming distance between t_x and f_x is equal to that between t_y and f_y. For the Euclidean distance, consider the Euclidean distance between x and y, which is equal to $\left(\sqrt{a^2 - a + 1}\right)$. This means that the distance between the vague value with the most imprecise evidence and the vague value with the most precise evidence is not equal to 1. (Actually, the Euclidean distance in this case is in the interval $[\sqrt{3}/2, 1)$.) However, our intuition shows that the distance in this case should always be equal to 1.

In order to deal with these problems, Yao proposed the concept of the similarity measure between the vague values [9], by using projection and distance between two vague sets. Let $x = [t_x, 1 - f_x]$ and $y = [t_y, 1 - f_y]$ be two vague values, then projection of vague value x on vague value y is defined by

$$D_Y(x,y) = \frac{t_x t_y + (1 - f_x)(1 - f_y)}{\sqrt{(t_y)^2 + (1 - f_y)^2}}. \tag{11}$$

Distance between two vague sets is defined by

$$J_Y(x,y) = \sqrt{(t_x - t_y)^2 + (f_x - f_y)^2}. \tag{12}$$

From formulas (11) and (12), the comprehensive evaluation of similarity measure between vague values x and y can be expressed as:

$$M_Y(x,y) = \frac{D_Y(x,y)}{D_Y(x,y) + J_Y(x,y)}. \tag{13}$$

However, for example, let $D_Y(x,y) = \alpha J_Y(x,y)$, $\alpha \neq 0$ and $\alpha \neq 1$, we can obtain $M_Y(x,y) = \alpha/(\alpha + 1)$ from formula (13). This means that the similarity measure between difference x and one y is equal.

4 A New Similarity Measure

In order to solve all the problems mentioned above, we define a new similarity measure between the vague values. Let x and y be two vague values with respect to $u \in U$ in a universe of discourse U, where $x = [t_x(u), 1 - f_x(u)]$ and $y = [t_y(u), 1 - f_y(u)]$. We put the $t_x(u)$ and $f_x(u)$ as two sample values of an skillful test with respect to u, and the $t_y(u)$ and $f_y(u)$ as other two sample values with respect to u. Normally, for two sample values of an skillful test, we should consider the internal difference between two sample values, i.e.$S(x) = t_x(u) - f_x(u)$ (or $S(y) = t_y(u) - f_y(u)$), where $S(x)$ (or $S(y)$) is also called the kernel function. For two vague values $[t_x(u), 1 - f_x(u)]$ and $[t_y(u), 1 - f_y(u)]$, the comparison of internal difference of $S(x)$ and $S(y)$ is defined as follows:

$$M_I(x, y) = \frac{1 + \min\{(t_x(u) - f_x(u)), (t_y(u) - f_y(u))\}}{1 + \max\{(t_x(u) - f_x(u)), (t_y(u) - f_y(u))\}}. \tag{14}$$

Considering the internal difference of two sample values at the same time, we consider the external difference of a pair of samples and another pair of samples. Based on the set-theoretic approach, Pappis and Karacapilidis defined the similarity measure between fuzzy sets A and B as follows [10]:

$$T_P(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{\sum_{i=1}^{n} (a_i \wedge b_i)}{\sum_{i=1}^{n} (a_i \vee b_i)}. \tag{15}$$

It is extensively accepted that the vague sets are regarded as the extended fuzzy sets. In other words, the fuzzy sets are a special case of vague sets. So we can define the similarity measure between vague values based on the set-theoretic approach. Let $x = [t_x(u), 1 - f_x(u)]$ and $y = [t_y(u), 1 - f_y(u)]$ be the two vague values. Following on the similarity measure of Eq. (15), the degree of similarity between the vague values x and y can be evaluated by the function $M_E(x, y)$ as follows:

$$\begin{aligned} M_E(x, y) &= \frac{t_x(u) \wedge t_y(u) + (1 - f_x(u)) \wedge \left(1 - f_y(u)\right)}{t_x(u) \vee t_y(u) + (1 - f_x(u)) \vee \left(1 - f_y(u)\right)} \\ &= \frac{\min\{(t_x(u), t_y(u)\} + \min\{(1 - f_x(u)), (1 - f_y(u))\}}{\max\{(t_x(u), t_y(u)\} + \max\{(1 - f_x(u)), (1 - f_y(u))\}}. \end{aligned} \tag{16}$$

According to the analysis of internal difference and external difference of two vague values, the following definition of similarity measure between the vague values x and y.

Definition 9 Let x and y be two vague values with respect to $u \in U$ in a universe of discourse U, where $x = [t_x(u), 1 - f_x(u)]$ and $y = [t_y(u), 1 - f_y(u)]$, the similarity measure between the vague values x and y is defined by

$$M(x,y)=\frac{1}{2}(M_I(x,y)+M_E(x,y))=\frac{1}{2}\left(\frac{1+\min\{(t_x(u)-f_x(u)),(t_y(u)-f_y(u))\}}{1+\max\{(t_x(u)-f_x(u)),(t_y(u)-f_y(u))\}}+\frac{\min\{(t_x(u),t_y(u)\}+\min\{(1-f_x(u)),(1-f_y(u))\}}{\max\{(t_x(u),t_y(u)\}+\max\{(1-f_x(u)),(1-f_y(u))\}}\right) \tag{17}$$

The large the value of $M(x,y)$, the more the similarity between the vague values x and y. From Definition 9, we can obtain the following theorem.

Theorem 1 The following statements are true:

(1) The similarity degree is bounded, i.e. $0 \le M(x,y) \le 1$.
(2) If the vague values x and y are equal, then $M(x,y)=1$.
(3) The similarity measure is commutative, i.e. $M(x,y) = M(y,x)$.
(4) The similarity degree between the zero vague value and any non-zero vague value is equal to 0.
(5) If $x=[1,1]$ and $y=[0,1]$, then $M(x,y)=0.5$.
(6) Let $z=[t_z(u),1-f_z(u)]$, if $x\subseteq y\subseteq z$, then $M(x,z)\le M(x,y)$ and $M(x,z)\le M(y,z)$.

Proof (1)–(5) are obtained directly from the definition of the $M(x,y)$ model, respectively. We only prove statement (6). If $x\subseteq y\subseteq z$, then $t_x(u)\le t_y(u)\le t_z(u)$, $f_x(u)\ge f_y(u)\ge f_z(u)$.

So,

$$t_x(u)-f_x(u)+1\le t_y(u)-f_y(u)+1\le t_z(u)-f_z(u)+1.$$

Then

$$M(x,y)=\frac{1}{2}\left(\frac{1+t_x(u)-f_x(u)}{1+t_y(u)-f_y(u)}+\frac{t_x(u)+1-f_x(u)}{t_y(u)+1-f_y(u)}\right)=\frac{t_x(u)-f_x(u)+1}{t_y(u)-f_y(u)+1}.$$

In a similar way, we can get similarity measure $M(x,z)$ and similarity measure $M(y,z)$ as follows:

$$M(x,z)=\frac{1}{2}\left(\frac{1+t_x(u)-f_x(u)}{1+t_z(u)-f_z(u)}+\frac{t_x(u)+1-f_x(u)}{t_z(u)+1-f_z(u)}\right)=\frac{t_x(u)-f_x(u)+1}{t_z(u)-f_z(u)+1}$$

and

$$M(y,z)=\frac{1}{2}\left(\frac{1+t_y(u)-f_y(u)}{1+t_z(u)-f_z(u)}+\frac{t_y(u)+1-f_y(u)}{t_z(u)+1-f_z(u)}\right)=\frac{t_y(u)-f_y(u)+1}{t_z(u)-f_z(u)+1}.$$

Therefore,

$$M(x,z)=\frac{t_x(u)-f_x(u)+1}{t_z(u)-f_z(u)+1}\leq\frac{t_x(u)-f_x(u)+1}{t_y(u)-f_y(u)+1}=M(x,y),$$
$$M(x,z)=\frac{t_x(u)-f_x(u)+1}{t_z(u)-f_z(u)+1}\leq\frac{t_y(u)-f_y(u)+1}{t_z(u)-f_z(u)+1}=M(y,z).$$

Next we generalize the similarity measure to two given vague sets.

Definition 10 Let A and B be two vague sets in the universe of discourse $U=\{u_1,u_2\ldots,u_n\}, A=\sum_{i=1}^{n}[t_A(u_i),1-f_A(u_i)]/u_i, B=\sum_{i=1}^{n}[t_B(u_i),1-f_B(u_i)]/u_i$. The similarity measure between the vague sets A and B can be evaluated follows:

$$M(A,B)=\frac{1}{2n}\sum_{i=1}^{n}\left[\frac{1+\min\{(t_A(u_i)-f_A(u_i)),(t_B(u_i)-f_B(u_i))\}}{1+\max\{(t_A(u_i)-f_A(u_i)),(t_B(u_i)-f_B(u_i))\}}\right.$$
$$\left.+\frac{\min\{t_A(u_i),t_B(u_i)\}+\min\{(1-f_A(u_i)),(1-f_B(u_i))\}}{\max\{t_A(u_i),t_B(u_i)\}+\max\{(1-f_A(u_i)),(1-f_B(u_i))\}}\right]. \quad (18)$$

Similarly, we give the definition of distance between two vague sets to set as $D(A,B)=1-M(A,B)$.

From Definition 10, we obtain the following theorem for vague sets, which is similar to Theorem 1.

Theorem 2 The following statements related to $M(A,B)$ are true:

(1) The similarity measure is bounded, i.e. $0\leq M(A,B)\leq 1$.
(2) $M(A,B)=1$, if and only if, the vague sets A and B are equal (i.e. $A=B$).
(3) $M(A,B)=0$, if and only if, all the vague sets $A=\sum_{i=1}^{n}[0,0]/u_i$ or $B=\sum_{i=1}^{n}[0,0]/u_i$.
(4) The similarity measure is commutative, i.e. sets $M(A,B)=M(B,A)$.

5 Numerical Experiments

We use the ideas of this paper for choosing the best supplier. There are five suppliers A_1, A_2, A_3, A_4, A_5 selected as alternatives, against three attributes C_1, C_2 and C_3. We obtain decision matrix A as follows:

$$A=\begin{matrix}A_1\\A_2\\A_3\\A_4\\A_5\end{matrix}\begin{bmatrix}(C_1,[0.2,0.8]) & (C_2,[0.3,0.9]) & (C_3,[0.2,0.9])\\(C_1,[0.3,0.7]) & (C_2,[0.2,0.8]) & (C_3,[0.3,0.8])\\(C_1,[0.5,0.6]) & (C_2,[0.5,0.6]) & (C_3,[0.4,0.7])\\(C_1,[0.5,0.7]) & (C_2,[0.4,0.6]) & (C_3,[0.5,0.7])\\(C_1,[0.4,0.6]) & (C_2,[0.6,0.7]) & (C_3,[0.6,0.6])\end{bmatrix}.$$

The paper takes the weight of each attribute as 1. The reference sequence A^* as follows is composed of the optimal interval value of indicator over all alternatives $A^* = [(C_1,\ [1,1]),\ (C_2, [1,1]),\ (C_3, [1,1])]$.

Next, we calculate the order vector of five suppliers. That is, to calculate the similarity measures between two vague sets (denoted as A_1, A_2, A_3, A_4, A_5 and A^*), respectively. According to formula (18), we can obtain the order vector as follows:

$$\begin{aligned}&(M(A_1, A^*), M(A_2, A^*), M(A_3, A^*), M(A_4, A^*), M(A_5, A^*))\\&= (0.550, 0.517, 0.550, 0.567, 0.583).\end{aligned}$$

Therefore, the ranking order of five suppliers will be as follows:

$$A_5 \succ A_4 \succ A_3 \sim A_1 \succ A_2.$$

We can say that supplier A_5 is the best supplier among the five suppliers.

6 Conclusion

Many similarity measures have been proposed in literature for measuring the degree of similarity between fuzzy sets. Also several efforts have been made for the similarity measure between vague sets. After analyzing the limitations in current similarity measures for vague sets, we have proposed a new method for measuring the similarity between vague sets in this paper. It is the basic method between vague sets to deeply understand the support, the difference of true-membership and the difference of false-membership, to significantly distinguish the directions of difference (positive and negative).

Acknowledgements The research was supported by the National Natural Science Foundation of China (Grant No. 60972115) and the Scientific Research Common Program of Beijing Municipal Commission of Education (Grant No. SQKM201211232016).

References

1. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
2. Gau, W.L., Buehrer, D.J.: Vague sets. IEEE Trans. Syst. Man Cybern. (Part B) **23**(2), 610–614 (1993)
3. Chen, S.M.: Measures of similarity between vague sets. Fuzzy Sets Syst. **74**(2), 217–223 (1995)
4. Hong, D.H., Kim, C.: A note on similarity measures between vague sets and elements. Inf. Sci. **115**, 83–96 (1999)
5. Li, F., Xu, Z.Y.: Similarity measures between vague sets. Chin. J. Softw. **12**(6), 922–927 (2001)

6. Li, Y.H., Chi, Z.X., Yan, D.Q.: Vague similarity and vague entropy. Comput. Sci. (Chin. J.) **29**(12), 129–132 (2002)
7. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst. **114**, 505–518 (2000)
8. Grzegorzewski, P.: Distances between intuitionistic fuzzy sets and/or interval valued fuzzy sets based on the Hausdorff metric. Fuzzy Sets Syst. **148**, 319–328 (2004)
9. Yao, R.P., Shen, H.Z.: Fuzzy multicriteria decision making based on projectivity and distance of vague sets. J. Math. Pract. Theory **32**(2), 19–22 (2009)
10. Pappis, C.P., Karacapilidis, N.I.: A comparative assessment of measures of similarity of fuzzy values. Fuzzy Sets Syst. **56**, 171–174 (1993)
11. Zhang, W.N., Wang, K.: An efficient evaluation of a fuzzy equi-join using fuzzy equality indicators. IEEE Trans. Knowl. Data Eng. **12**(2), 225–237 (2000)
12. Ma, Z.M., Zhang, W.J., Ma, W.Y.: Semantic measure of fuzzy data in extended possibility-based fuzzy relational database. Int. J. Intell. Syst. **15**(8), 705–716 (2000)
13. Lu, A., Ng, W.: Managing merged data by vague functional dependencies. Lect. Notes Comput. Sci. **3288**, 259–272 (2004)
14. Li, Y.H., David, L.Olson, Qin, Z.: Similarity measures between intuitionistic fuzzy (vague) sets: A comparative analysis. Pattern Recogn. Lett. **28**(2), 278–285 (2007)
15. Zhang, D.F., Zhang, J.L., Lai, K.K., Lu, Y.: An novel approach to supplier selection based on vague sets group decision. Expert Syst. Appl. **36**(5), 9557–9563 (2009)
16. Zhang, Q.S., Jiang, S.Y.: A note on information entropy measures for vague sets and its applications. Inf. Sci. **178**, 4184–4191 (2008)
17. Ye, J.: Using an improved measure function of vague sets for multicriteria fuzzy decision-making. Expert Syst. Appl. **37**, 4706–4709 (2010)

Part VII
Others

The Resolution Method for 10-Element Linguistic Truth-Valued Propositional Logic

Li Zou, Ying Wang, Ying-ying Xu and Min-wen Yang

Abstract This paper analyzes five kinds of complementary literals based on the 10-element linguistic truth-valued propositional logic system. In order to give an effective resolution method, we define a linguistic hedges compound operation $\oplus$ and study complementary of linguistic truth-values. Then the resolution method based on a filter J of the system and the corresponding algorithm is constructed. An example is given to illustrate the validity of the proposed method.

Keywords Lattice implication algebra · Linguistic truth-valued propositional logic · Resolution principle · Complementary literals

1 Introduction

People express themselves through natural languages in the real life. Natural languages can be formalized by L.A. Zadeh's linguistic variables [1]. The truth-valued proposition is not often exactly true or false, however accompany with linguistic hedges, such as absolutely, highly, very, quite, exactly, almost, rather, somewhat, slightly and so on. A linguistic hedge possesses an intuitive meaning which can be expressed by a semantic ordering relation, e.g. *ApproximatelyTrue* < *True*, while *VeryTrue* > *True*. These truth values are called linguistic truth values [2, 3].

Since 1965, automated reasoning based on Robinson's [4] resolution rule has been extensively studied in the context of proving the system [5–7] to support a wide

L. Zou (✉) · Y. Wang · Y. Xu
School of Computer and Information Technology, Liaoning Normal University, Dalian 116081, China
e-mail: zoulicn@163.com

Y. Wang
e-mail: wangying1755@126.com

M. Yang
Senior Three Class Eight, The Affiliated High School of Peking University, Beijing 100080, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_56

spectrum of computational tasks. They are widely applied to areas such as artificial intelligence, logic programming, problem solving and question answering systems, database theory, and so on.

Lattice-valued logic system which has been presented by Xu [8, 9] is an important case of multi-valued logic. Since 1993, lattice implication algebra (LIA) and some related discussion on logic system based on LIA have been invested. Nowadays, there exist many alternative methods for linguistic values based lattice implication algebra. Such as X.X. He got the-unit algorithm in $L_nP(X)$ [10]. L. Zou and P. Shi et al. have proposed a kind of linguistic truth-valued intuitionistic fuzzy lattice based on the view of intuitionistic fuzzy set and linguistic truth-valued lattice implication algebra [11] and so on. In fact, there are some linguistic truth-valued are incomparable. The research of resolution principle and algorithm based on LIA will help the realization on intelligent system which includes incomparable elements. About the resolution algorithm, there is an obvious fact that, the more limit and the less clauses, the more simple for the due process thereby. In this respect, X.M. Zhong et al. studied on quasi-lock semantic resolution method for linguistic truth-valued lattice-valued propositional logic $L_{V(n\times 2)}P(X)$ [12].

This paper is organized as follows: Sect. 2 reviews some preliminary relevant concepts and 10-element linguistic truth-valued propositional logic 10LTVP is established. In Sect. 3, resolution method for 10LTVP is studied. Finally, some concluding remarks will be given in Sect. 4.

2 10-Element Linguistic Truth-Valued Propositional Logic

2.1 Lattice Implication Algebra

Definition 1 [13] Let $(L, \wedge, \vee, ', I, O)$ be a bounded lattice with an order-reversing involution "$'$", I and O are the greatest and the least element of L respectively. For any $x, y, z \in L$, if mapping $\rightarrow: L \times L \rightarrow L$ satisfies:

(1) $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$;
(2) $x \rightarrow x = I$;
(3) $x \rightarrow y = y' \rightarrow x'$;
(4) $x \rightarrow y = y \rightarrow x$, implies $x = y$;
(5) $(x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x$;
(6) $(x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z)$;
(7) $(x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z)$.

Then $(L, \wedge, \vee, ', I, O)$ is called a lattice implication algebra.

Definition 2 [12] Let $AD_n = \{c_1, c_2, \ldots, c_n\}$ be a set with n modifiers and $c_1 < c_2 < \cdots < c_n$, $MT = \{f, t\}$ be a set of meta truth values, and $f < t$. Denote $L_{V(n\times 2)} = AD_n \times MT$.

Define a mapping g as $g : L_{V(n\times 2)} \to L_n \times L_2$,

$$g((c_i, mt)) = \begin{cases} (a_i, b_1), mt = f \\ (a_i, b_2), mt = t, \end{cases}$$

then g is a bijection, denote its inverse mapping as g^{-1}. For any x, $y \in L_{V(n\times 2)}$, define $x \vee y = g^{-1}(g(x) \vee g(y))$, $x \wedge y = g^{-1}(g(x) \wedge g(y))$, $x' = g^{-1}((g(x))')$, $x \to y = g^{-1}(g(x) \to g(y))$ then $L_{V(n\times 2)} = (L_{V(n\times 2)}, \vee, \wedge, ', \to, (c_n, f), (c_n, t))$ is called a linguistic truth-valued lattice implication algebra generated by AD_n and MT.

2.2 Related Concepts on 10-Element Linguistic Truth-Valued Propositional Logic

Definition 3 Let $H = \{h_i | i = 0, 1, 2, 3, 4\}$ be the linguistic hedge operator set where $h_0 = slightly$, $h_1 = somewhat$, $h_2 = rather$, $h_3 = very$, $h_4 = absolutely$. H has a natural order structure, i.e. $h_0 < h_1 < h_2 < h_3 < h_4$. '$\oplus$' is a dyadic operator of H which satisfies with Table 1:

$$h_i \oplus h_j = \begin{cases} h_{\min\{4, i-2+j\}}, i + j \geq 2 \\ h_0, i + j < 2. \end{cases}$$

Proposition 1 *For any h, h_i, $h_j \in H$,*

(1) $h \oplus \max\{h_i, h_j\} = \max\{h \oplus h_i, h \oplus h_j\}$;
(2) $h \oplus \min\{h_i, h_j\} = \min\{h \oplus h_i, h \oplus h_j\}$.

Proof 1. When $h = h_0$, according to Definition 3, we have
$h_0 \oplus \max\{h_0, h_1\} = h_0 \oplus h_1 = h_0 = \max\{h_0 \oplus h_0, h_0 \oplus h_1\}$,
$h_0 \oplus \max\{h_1, h_2\} = h_0 \oplus h_2 = h_0 = \max\{h_0 \oplus h_1, h_0 \oplus h_2\}$,
$h_0 \oplus \max\{h_2, h_3\} = h_0 \oplus h_3 = h_1 = \max\{h_0 \oplus h_2, h_0 \oplus h_3\}$,
$h_0 \oplus \max\{h_3, h_4\} = h_0 \oplus h_4 = h_2 = \max\{h_0 \oplus h_3, h_0 \oplus h_4\}$;

Table 1 Operation $\oplus$

$\oplus$	h_4	h_3	h_2	h_1	h_0
h_4	h_4	h_4	h_4	h_3	h_2
h_3	h_4	h_4	h_3	h_2	h_1
h_2	h_4	h_3	h_2	h_1	h_0
h_1	h_3	h_2	h_1	h_0	h_0
h_0	h_2	h_1	h_0	h_0	h_0

2. When $h = h_1$, according to Definition 3, we have
   $h_1 \oplus \max\{h_0, h_1\} = h_1 \oplus h_1 = h_0 = \max\{h_1 \oplus h_0, h_1 \oplus h_1\}$,
   $h_1 \oplus \max\{h_1, h_2\} = h_1 \oplus h_2 = h_1 = \max\{h_1 \oplus h_1, h_1 \oplus h_2\}$,
   $h_1 \oplus \max\{h_2, h_3\} = h_1 \oplus h_3 = h_2 = \max\{h_1 \oplus h_2, h_1 \oplus h_3\}$,
   $h_1 \oplus \max\{h_3, h_4\} = h_1 \oplus h_4 = h_3 = \max\{h_1 \oplus h_3, h_1 \oplus h_4\}$;
3. When $h = h_2$, according to Definition 3, we have
   $h_2 \oplus \max\{h_0, h_1\} = h_2 \oplus h_1 = h_1 = \max\{h_2 \oplus h_0, h_2 \oplus h_1\}$,
   $h_2 \oplus \max\{h_1, h_2\} = h_2 \oplus h_2 = h_2 = \max\{h_2 \oplus h_1, h_2 \oplus h_2\}$,
   $h_2 \oplus \max\{h_2, h_3\} = h_2 \oplus h_3 = h_3 = \max\{h_2 \oplus h_2, h_2 \oplus h_3\}$,
   $h_2 \oplus \max\{h_3, h_4\} = h_2 \oplus h_4 = h_4 = \max\{h_2 \oplus h_3, h_2 \oplus h_4\}$.

The same goes for h_3, h_4. To sum up, for any h, h_i, $h_j \in H$, $h \oplus \max\{h_i, h_j\} = \max\{h \oplus h_i, h \oplus h_j\}$ is found. The same can be proved in (2). So the operation $\oplus$ can limit the modified multiple linguistic truth-valued in a specific linguistic truth-valued lattice implication algebra.

Definition 4 Let $H = \{h_0, h_1, h_2, h_3, h_4\}$ and the basic truth value set $C = \{T, F\}$, where $T = true$, $F = false$. Let V be a linguistic truth value set, every linguistic truth value $v \in V$ is composed of a linguistic hedge operator and a basic word c, i.e. $V = H \times C$ where the linguistic hedge operator set H is linear and finite set. Denote $V = \{(h_4, T), (h_3, T), (h_2, T), (h_1, T), (h_0, T), (h_0, F), (h_1, F), (h_2, F), (h_3, F), (h_4, F)\}$, $L_{10} = (V, \vee, \wedge, \rightarrow)$, the operation "$\vee$" and "$\wedge$" the *Hasse* diagram of L_{10}, Fig. 1. $h_i, T' = h_i, F$ and its operations "$\rightarrow$" are defined by the following (defined by the Tables 1 and 2):

(1) $(h_i, T) \rightarrow (h_j, F) = (h_{\max\{0, i+j-4\}}, F)$;
(2) $(h_i, F) \rightarrow (h_j, T) = (h_{\min\{4, i+j\}}, T)$;
(3) $(h_i, T) \rightarrow (h_j, T) = (h_{\min\{4, 4-i+j\}}, T)$;
(4) $(h_i, F) \rightarrow (h_j, T) = (h_{\min\{4, 4-i+j\}}, T)$.

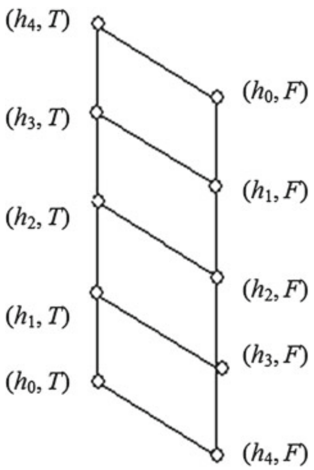


Fig. 1 *Hasse* diagram of L_{10}

Table 2 Implication operator of $L_{10} = (V, \vee, \wedge, ', \rightarrow)$

$\rightarrow$	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)	(h_0,T)	(h_4,F)	(h_3,F)	(h_2,F)	(h_1,F)	(h_0,F)
(h_4,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)	(h_0,T)	(h_4,F)	(h_3,F)	(h_2,F)	(h_1,F)	(h_0,F)
(h_3,T)	(h_4,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)	(h_3,F)	(h_2,F)	(h_1,F)	(h_0,F)	(h_0,F)
(h_2,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_2,F)	(h_1,F)	(h_0,F)	(h_0,F)	(h_0,F)
(h_1,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_3,T)	(h_1,F)	(h_0,F)	(h_0,F)	(h_0,F)	(h_0,F)
(h_0,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_0,F)	(h_0,F)	(h_0,F)	(h_0,F)	(h_0,F)
(h_4,F)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)
(h_3,F)	(h_4,T)	(h_4,T)	(h_4,T)	(h_4,T)	(h_3,T)	(h_3,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)
(h_2,F)	(h_4,T)	(h_4,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_2,T)	(h_3,T)	(h_4,T)	(h_3,T)	(h_2,T)
(h_1,F)	(h_4,T)	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)	(h_1,T)	(h_2,T)	(h_3,T)	(h_4,T)	(h_3,T)
(h_0,F)	(h_4,T)	(h_3,T)	(h_2,T)	(h_1,T)	(h_0,T)	(h_0,T)	(h_1,T)	(h_2,T)	(h_3,T)	(h_4,T)

So the $L_{10} = (V, \vee, \wedge, ', \rightarrow)$ to be the lattice implication algebra.

Definition 5 Let P be an atomic proposition, $h_i \in H$, then called $h_i P$ is an atomic proposition of 10*LTVP*.

For example: To make an atomic proposition P: Tom is lively, let $h_3 = very$, then the $h_3 P$ express "Tom is very lively" is an atomic proposition of 10*LTVP*.

Definition 6 The formula in 10*LTVP* is defined by the following recursive:

(1) 10*LTVP* atom hP is a formula;
(2) If A is a formula, then (hA), $\neg A$ is the formula, $h \in H$;
(3) If A, B is a formula, then to A, $(A \vee B)$, $(A \wedge B)$, $(A \rightarrow B)$ are formulas;
(4) All the formulas are the symbolic strings which use (1)–(3) with finite times.

Definition 7 A valuation of 10*LTVP* is a propositional algebra homomorphism, V : 10*LTVP* $\rightarrow L_{10}$.

Proposition 2 *For any* $P, Q \in$ 10*LTVP*, $v \in L_{10}$, *Have the following properties:*

(1) $\neg(vP) \equiv v(\neg P)$;
(2) $vP \equiv v'(\neg P)$;
(3) $v(A \vee B) \equiv vA \vee vB$;
(4) $v(A \wedge B) \equiv vA \wedge vB$.

Theorem 1 $J = \{(h_4,T), (h_3,T), (h_2,T), (h_1,T), (h_0,T)\}$ *is a filter in the* L_{10}, $u, v \in S$, *uP and vQ are the formula in* 10*LTVP*, $uP \rightarrow vQ \geq \neg uP \vee vQ$.

Proof When $u = (h_4,T)$, no matter what value of v, from Table 2 there always goes $uP \rightarrow vQ = \neg uP \vee vQ$. As the same, when $v = (h_4,T)$, no matter what value of u, there always goes $uP \rightarrow vQ = \neg uP \vee vQ$. When u, v are other elements in the J from Table 2 there always goes $uP \rightarrow vQ > \neg uP \vee vQ$. So there goes $uP \rightarrow vQ \geq \neg uP \vee vQ$, the definition is proved.

3 Resolution Method in 10*LTVP*

3.1 Algorithm of Transforming Any Formula to a Reducible Form

In order to use the resolution method, the formula should be transformed to a reducible form, which is absolutely necessary before resolution. There are a lot of presented results for classical logic and multi-valued logic based on *Kleene*'s implication to be found. However, for any formula F in 10*LTVP* based on *LIA*, implication connective can't be expressed by other connectives. So the procedure of transforming any formula to a normal form is different from that of the presented results, it must be solved firstly. In this section, we will discuss the transformation. It depends on different structure of formula and we will deal with it respectively:

Definition 8 Let $P \in 10LTVP$, $J = \{(h_4, T), (h_3, T), (h_2, T), (h_1, T), (h_0, T)\}$ be a filter in the 10*LTVP*. If there is an interpretation of C such that $C(P) \in J$, said the formula P is $J-$ *satisfiable*; if for any interpretation of C, $C(P) \in J$, said the formula P is $J-$ *true*; if for any interpretation of C, $C(P) \in J$, said P is $J-$ *false* formula.

Theorem 2 *For any* P, $Q \in 10LTVP$, $v, u \in J$, *if the* $u = (h_4, T)$, *there has* $uP \to vQ = \neg uP \vee vQ$; *if the* $u = (h_3, T)$, *there has* $uP \to vQ = (h_3) \oplus (\neg uP \vee vQ)$; *if the* $u = (h_1, T)$, *there has* $uP \to vQ = (h_3) \oplus (\neg uP \vee vQ)$; *if the u is other element, there has* $uP \to vQ = (h_4) \oplus (\neg uP \vee vQ)$.

Proof It can be proved by Proposition 1 and Theorem 1.

For example: The same for (h_4, T), (h_3, T), (h_1, T) and (h_0, T).

3.2 Complementary Literals

Definition 9 For any h_i, $h_j \in H$, u, $v \in J$, let uP and $v(\neg P)$ be the two generalized literals of 10*LTVP*, if there is any assignment $((uP)v(\neg P)) \in S$, then uP and $v(\neg P)$ are called $J-$ *complementary literals*.

A	B	$\to$	$h_4 \oplus (\neg A \vee B)$
(h_2, T)	(h_4, T)	(h_4, T)	(h_4, T)
	(h_3, T)	(h_4, T)	(h_4, T)
	(h_2, T)	(h_4, T)	(h_4, T)
	(h_1, T)	(h_3, T)	(h_3, T)
	(h_0, T)	(h_2, T)	(h_2, T)

(1) If $i = j$, then $((uP)v(\neg P)) = (h_4, T) \in J$, uP and $v(\neg P)$ are called absolutely strong complementary literals;
(2) If $|i - j| = 1$, then uP and $v(\neg P)$ are called very strong complementary literals;
(3) If $|i - j| = 2$, then uP and $v(\neg P)$ are called rather strong complementary literals;
(4) If $|i - j| = 3$, then uP and $v(\neg P)$ are called somewhat strong complementary literals;
(5) If $|i - j| = 4$, then uP and $v(\neg P)$ are called slightly strong complementary literals.

So the complementary degrees based on the five kinds of the complementary literals are defined.

Definition 10 (*Soft resolution principle*) Let C_1 and C_2 be the two clauses of 10 *LTVP*, uP and $v(\neg P)$ are disjunctive of C_1 and C_2, where uP and $v(\neg P)$ is $J-$ *complementary literals*. Delete the uP and $v(\neg P)$ from C_1 and C_2, then get the C_{r1} and C_{r2}, $C_{r1} \vee C_{r2}$ is *a* $J -$ *resolvent* of C_1 and C_2. The same:

(1) If the uP and $v(\neg P)$ are absolutely strong complementary literals, $C_{r1} \vee C_{r2}$ is C_1 and C_2 absolutely strong resolution, denoted as $(h_4, T)(C_{r1} \vee C_{r2})$;
(2) If the uP and $v(\neg P)$ are very strong complementary literals, $C_{r1} \vee C_{r2}$ is C_1 and C_2 very strong resolution, denoted as $(h_3, T)(C_{r1} \vee C_{r2})$;
(3) If the uP and $v(\neg P)$ are rather strong complementary literals, $C_{r1} \vee C_{r2}$ is C_1 and C_2 rather strong resolution, denoted as $(h_2, T)(C_{r1} \vee C_{r2})$;
(4) If the uP and $v(\neg P)$ are somewhat strong complementary literals, $C_{r1} \vee C_{r2}$ is C_1 and C_2 somewhat strong resolution, denoted as $(h_1, T)(C_{r1} \vee C_{r2})$;
(5) If the uP and $v(\neg P)$ are slightly strong complementary literals, $C_{r1} \vee C_{r2}$ is C_1 and C_2 slightly strong resolution, denoted as $(h_0, T)(C_{r1} \vee C_{r2})$.

3.3 Resolution Method

Theorem 3 *Let* $P \in 10PTVP$ *be a formula, P is the* $J - false$ *if and only if there exists a deduction from P which can deduce* $J -$ *null clause (denoted by* $J - \diamond$*).*

Proof From the above definition can be used in the proof process is attributed words and negative words. Completeness of logic resolution principle consists of two-valued logic. P is the $J - false$ if and only if exists a deduction from P which can deduce $J - \diamond$ clause.

In the soft resolution, resolution based on $J -$ *complementary literals*, only need to calculate the modal variables. Because mood variables are limited that can reach $J - nil$. Resolution algorithm 10*LTVP* is given below:

Step 1: Transforming formula to a reducible form;
Step 2: Finding out the truth values which are not in the J proposition. If the proposition is not in J, then consider the converse of this proposition, i.e. if there are

$(h_4, F)P$, $(h_3, F)P$, $(h_2, F)P$, $(h_1, F)P$, $(h_0, F)P$, will be converted into $(h_4, T)\neg P$, $(h_3, T)\neg P$, $(h_2, T)\neg P$, $(h_1, T)\neg P$, $(h_0, T)\neg P$, make its truth values in J;

Step 3: If there are $J-$*false clause* theorem of card, stop, otherwise go to step 4;

Step 4: Deleting the $J-$ *complementary literals*. The linguistic hedge operator of the resolvent comes from the qualitative operation of the linguistic hedge operator before the two resolution clauses;

Step 5: If the $J-$ *empty clause*, then theorem is proved, stop, otherwise go to step 4.

3.4 The Deduce Based on Resolution

Example: Prove that the formula $A \to C$ can be launched by the $A \to B$ and $B \to C$, and $V(A) = (h_4, F)$, $V(B) = (h_3, T)$, $V(C) = (h_2, T)$.

Proof the clause can be obtained by the known:

1. $(h_4, F)\neg A(h_3, T)B = (h_4, T)A \vee (h_3, T)B$
2. $(h_3) \oplus [(h_3, T)\neg B \vee (h_2, T)C] = (h_4, T)\neg B \vee (h_3, T)C$
3. $(h_4, F)A = (h_4, T)\neg A$
4. $(h_2, T)\neg C$
5. $[(h_4, T) \oplus (h_3, T)]B = (h_4, T)B$ 1 and 3
6. $[(h_3, T) \oplus (h_4, T)]B = (h_4, T)\neg B$ 2 and 4
7. $(h_4, T) - nil$ 5 and 6

So the conclusion by $A \to B$ and $B \to C$:$A \to C$. This theorem is due to the linguistic hedge operator is absolutely strong resolution, that is, the theorem is established.

4 Conclusion

People use the qualitative model to deal with vague or imprecise knowledge rather than quantitative model. We often use some linguistic values instead of numbers to think then getting conclusions. In this paper, we establish the 10-element linguistic truth-valued logic system which can express both comparable and incomparable information. In the resolution process linguistic truth-valued directly involved in the operation, the results also with the linguistic values. Automatic reasoning based on linguistic truth value attributed to artificial intelligence not only provides a new reasoning algorithm, but also can be used in various fields of decision system, risk analysis system and expert system.

Acknowledgments This work is partially supported by National Natural Science Foundation of China (Nos. 61372187 and 61173100) and National Natural Science Foundation of Liaoning Province (No. 2015020059).

References

1. Zadeh, L.A.: Toward a generalized theory of uncertainty (GTU)-an outline. Inf. Sci. **172**, 1–40 (2005)
2. Meng, D., Xu, Y., Qiu, X.P., Qin, K.Y.: Resolution principle based on six lattice-valued proposition logic LP'(X). In: Proceedings of Second International Conference on Machine and Cybernetics (2003)
3. Zou, L., Fang, S., Xu, Y.: Resolution method of six-element linguistic truth-valued intuitionistic propositional logic. In: 3rd International Conference on Proceedings of 2008 Intelligent System and Knowledge Engineering, pp. 141–145 (2008)
4. Robinson, J.P.: A Machine-oriented logic based on the resolution principle. J. ACM **12**, 23–41 (1965)
5. Zhang, H.L., Xu, B.Q., Zou, L.: An approach for knowledge representation based on ten-element lattice implication algebra. Comput. Appl. Softw. **30**(1), 37–40 (2013)
6. Xu, W.T., Zhang, W.Q., Xu, Y., Zhang, D.X.: Generalized resolution principle based on the lattice-valued first-order logic system. J. Xidian Univ. **41**(1), 168–173 (2014)
7. Zou, L., Tan, X.W., Zhang, Y.X.: Knowledge reasoning based on linguistic truth-valued intuitionstic fuzzy logic. Comput. Sci. **41**(1), 134–137 (2014)
8. Xu, Y., Ruan, D., Kerre, E.E., Jun, L.: CL-resolution principle based on first-order lattice-valued logic LF(X). Inf. Sci. **132**, 221–239 (2001)
9. Li, X.B., Xu, Y.: Automated reasoning algorithm based on lattice-valued first-order logic LF(x). Comput. Eng. Appl. **46**(23), 18–49 (2010)
10. He, X.X., Xu, Y., Liu, J.: Input and unit resolution methods for generalized horn clause set. J. Donghua Univ. (Eng. Ed.) **29**(1), 66–70 (2012)
11. Zou, L., Shi, P., Pei, Z., Xu, Y.: On an algebra of linguistic truth-valued intuitionistic lattice-valued logic. J. Intell. Fuzzy Syst. **24**, 447–456 (2013)
12. Zhong, X.M., Xu, Y., Liu, J., Ruan, D., Chen, S.W.: General form of resolution based on linguistic truth-valued lattice-valued logic. Soft Comput. **10**, 1767–1781 (2012)
13. Meng, D., Xu, Y., Qin, K.Y.: Algorithm of transforming any formula in LP(X) to a reducible form. In: Fuzzy Information Process Society, 2002 Annaul Meeting of the North American

Existence of R_0 Type Reverse Triple I Solutions

Jian Hou and Fei Yuan

Abstract In this paper, some examples are given to show that eight kinds of R_0 type reverse triple I solutions do not exist. Here considered are R_0 type reverse triple I FMP and FMT sustaining methods, R_0 type α-reverse triple I FMP and FMT sustaining methods, R_0 type reverse triple I FMP and FMT restriction methods and R_0 type α-reverse triple I FMP and FMT restriction methods.

Keywords Triple I method · Reverse triple I method · R_0 type reverse triple I method · Reverse triple I solution

1 Introduction

In 1973, Zadeh introduced the compositional rule of fuzzy inference (CRI) [1]. Afterwards, various methods of fuzzy reasoning have been proposed and a lot of successful applications in fuzzy system have been carried out [2–4]. Nowadays, fuzzy reasoning has become a theoretic foundation and an important tool for design and analysis of fuzzy controllers. In fuzzy reasoning, the following inference forms of Fuzzy Modus Ponens (FMP) and Fuzzy Modus Tollens (FMT) are often used:

$$\text{FMP}: \frac{\begin{array}{ll}\text{rule base} & A \longrightarrow B \\ \text{for given} & A^*\end{array}}{\begin{array}{ll}\text{to determine} & \qquad B^*\end{array}} \qquad \text{FMT}: \frac{\begin{array}{ll}\text{rule base} & A \longrightarrow B \\ \text{for given} & \qquad\quad B^*\end{array}}{\begin{array}{ll}\text{to determine} & A^*\end{array}}$$

where A and A^* are fuzzy sets in the domain X, and B and B^* are fuzzy sets in the domain Y. CRI has been widely applied successfully in various fields of industrial control. However, in view of logic semantics, there exist several problems. To

J. Hou (✉) · F. Yuan
Department of Applied Mathematics, Beijing Normal University,
Zhuhai 519087, China
e-mail: houjian@bnuz.edu.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_57

improve the CRI algorithm, the triple I method with total inference rules of fuzzy reasoning was proposed by Wang [5, 6]. The principle of triple I FMP is: suppose that X and Y are nonempty sets, $A, A^* \in \mathcal{F}(X)$ and $B \in \mathcal{F}(Y)$, then $B^* \in \mathcal{F}(Y)$ is the minimal fuzzy set such that

$$(A(x) \longrightarrow B(y)) \longrightarrow (A^*(x) \longrightarrow B^*(y)) \quad (1)$$

takes its maximum for any $x \in X$ and any $y \in Y$, where $\mathcal{F}(X)$ and $\mathcal{F}(Y)$ denote, respectively, the collections of all fuzzy subsets of X and Y. B^* is called the triple I FMP solution of (1). The principle of triple I FMT is as follows: Suppose that X and Y are nonempty sets, $A \in \mathcal{F}(X)$ and $B, B^* \in \mathcal{F}(Y)$, then $A^* \in \mathcal{F}(X)$ is the maximal fuzzy set such that (1) takes its maximum for any $x \in X$ and any $y \in Y$. And A^* is called the triple I FMT solution of (1). Based on the principle of triple I for fuzzy reasoning, Wang gave R_0 type reverse triple I FMP and FMT methods, where

$$R_0(a, b) = \begin{cases} 1, & a \le b \\ a' \vee b, & a > b \end{cases}, \quad a' = 1 - a.$$

Furthermore, Song proposed the principle of reverse triple I for fuzzy reasoning [7, 8].

In this paper, some examples are given to show that eight kinds of R_0 type reverse triple I solutions do not exist.

2 Existence of R_0 Type Reverse Triple I Solutions

2.1 *Existence of R_0 Type Reverse Triple I Sustaining Solutions*

In this section, we consider the reverse triple I sustaining methods. First, we consider the problem of reverse triple I FMP.

Sustaining principle of reverse triple I FMP [7] Suppose that X and Y are nonempty sets, $A, A^* \in \mathcal{F}(X)$ and $B \in \mathcal{F}(Y)$. Then $B^* \in \mathcal{F}(Y)$ satisfying this principle is the maximum fuzzy set such that

$$(A^*(x) \longrightarrow B^*(y)) \longrightarrow (A(x) \longrightarrow B(y)) \quad (2)$$

takes its maximum for any $x \in X$ and any $y \in Y$. B^* is called the reverse triple I FMP sustaining solution of (2).

Remark 1 Concerning the principle of reverse triple I FMP, when $A, A^* \in \mathcal{F}(X)$ and $B \in \mathcal{F}(Y)$ are known, there does not always exist the maximal fuzzy set B^* such that (2) takes its maximum for any $x \in X$ and any $y \in Y$.

Example 1 Suppose $X = Y = [0,1], A, A^* \in \mathcal{F}(X), B \in \mathcal{F}(Y), A(x) = \frac{2}{3}, A^*(x) = \frac{1}{2}$, $B(y) = \frac{1}{2}$. Then for any $x \in X$ and any $y \in Y$, there does not exist the maximal fuzzy set $B^* \in \mathcal{F}(Y)$ that makes (2) take its maximum. In fact, we have $(A^*(x))' = \frac{1}{2}$, $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$. Since the implication operator R_0 is monotonically increasing with respect to the first variable and is monotonically decreasing with respect to the second variable, the maximum of (2) is $R_0(R_0(A^*(x), 0), R_0(A(x), B(y))) = R_0(R_0(\frac{1}{2}, 0), R_0(\frac{2}{3}, \frac{1}{2})) = R_0(\frac{1}{2}, \frac{1}{2}) = 1$. However, (2) is expressed by

$$R_0(R_0(\frac{1}{2}, B^*(y)), R_0(\frac{2}{3}, \frac{1}{2})) = R_0(R_0(\frac{1}{2}, B^*(y)), \frac{1}{2}) = \begin{cases} 1, & R_0(\frac{1}{2}, B^*(y)) \le \frac{1}{2} \\ \frac{1}{2}, & R_0(\frac{1}{2}, B^*(y)) > \frac{1}{2} \end{cases}.$$

Thus, if (2) takes its maximum 1, then $R_0(\frac{1}{2}, B^*(y)) \le \frac{1}{2}$ should be satisfied. But

$$R_0(\frac{1}{2}, B^*(y)) = \begin{cases} 1, & \frac{1}{2} \le B^*(y) \\ \frac{1}{2}, & \frac{1}{2} > B^*(y) \end{cases} \ge \frac{1}{2}.$$

Obviously, only if $R_0(\frac{1}{2}, B^*(y)) = \frac{1}{2}$, (2) can take its maximum 1. Furthermore, by $R_0(\frac{1}{2}, B^*(y)) = \frac{1}{2}$, we get $B^*(y) < \frac{1}{2}$. Since $R_0(\frac{1}{2}, B^*(y))$ is not left continuous at $B^*(y) = \frac{1}{2}$, there does not exist the maximal $B^*(y)$ such that (2) takes its the maximum.

By the above example, we can see that R_0 type reverse triple I FMP sustaining solution of (2) does not always exist.

Similarly, we can consider the problem of reverse triple I FMT.

Sustaining principle of reverse triple I FMT [7]. Suppose that X and Y are nonempty, $A \in \mathcal{F}(X)$, $B, B^* \in \mathcal{F}(Y)$. Then $A^* \in \mathcal{F}(X)$ is the minimal fuzzy set that makes (2) take its maximum for any $x \in X$ and any $y \in Y$. And A^* is called reverse triple I FMT sustaining solution of (2).

Remark 2 Concerning the principle of reverse triple I FMT, for given $A \in \mathcal{F}(X)$ and $B, B^* \in \mathcal{F}(Y)$, there may be no the minimal fuzzy set A^* that makes (2) take its maximum for any $x \in X$ and any $y \in Y$, so there may be no the optimal solution.

Example 2 Suppose $X = Y = [0,1]$, $A(x) = \frac{2}{3}$, $B(y) = \frac{1}{2}$, $B^*(y) = \frac{1}{2}$. Then for any $x \in X$ and $y \in Y$, there does not exist the minimal fuzzy set A^* that makes (2) take its maximum. In fact, we have $(B^*(y))' = \frac{1}{2}$ and $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$. Since the implication operator R_0 is monotonically increasing with respect to the first variable, the maximum of (2) is $R_0(R_0(1, B^*(y)), R_0(A(x), B(y))) = R_0(B^*(y), R_0(\frac{2}{3}, \frac{1}{2})) = R_0(\frac{1}{2}, \frac{1}{2}) = 1$. Below, for given $A(x), B(y)$ and $B^*(y)$, we solve the minimal fuzzy set A^* that makes (2) take the maximum 1. By the definition of R_0, (2) is represented by

$$R_0(R_0(A^*(x), \frac{1}{2}), R_0(\frac{2}{3}, \frac{1}{2})) = R_0(R_0(A^*(x), \frac{1}{2}), \frac{1}{2}) = \begin{cases} 1, & R_0(A^*(x), \frac{1}{2}) \le \frac{1}{2} \\ \frac{1}{2}, & R_0(A^*(x), \frac{1}{2}) > \frac{1}{2} \end{cases},$$

If (2) takes the maximum 1, then $R_0(A^*(x), \frac{1}{2}) \le \frac{1}{2}$. By

$$R_0(A^*(x), \frac{1}{2}) = \begin{cases} 1, & A^*(x) \le \frac{1}{2} \\ \frac{1}{2}, & A^*(x) > \frac{1}{2} \end{cases},$$

we have $R_0(A^*(x), \frac{1}{2}) = \frac{1}{2}$. So $A^*(x) > \frac{1}{2}$. Since $R_0(A^*(x), \frac{1}{2})$ is not right continuous at $A^*(x) = \frac{1}{2}$, there does not exist the minimal fuzzy set A^* that makes (2) take its maximum.

By this example, we draw the conclusion that R_0 type reverse triple I FMT sustaining solution does not always exist.

2.2 Existence of R_0 Type α-Reverse Triple I Sustaining Solutions

This section is devoted to discussing existence of the generalization of R_0 type reverse triple I solutions. First, we consider the problem of reverse triple I FMP.

Sustaining principle of α-reverse triple I FMP [7]. Suppose that X and Y are nonempty, $A, A^* \in \mathcal{F}(X)$, $B \in \mathcal{F}(Y)$. For a given $\alpha \in (0,1]$, $B^* \in \mathcal{F}(Y)$ satisfying this principle is the maximal fuzzy set such that

$$(A^*(x) \longrightarrow B^*(y)) \longrightarrow (A(x) \longrightarrow B(y)) \geq \alpha \tag{3}$$

for any $x \in X$ and any $y \in Y$.

Similarly to the FMP problem, R_0 type α-reverse triple I FMP sustaining solution also does not always exist.

Example 3 Suppose that X, Y, A, B, A^* are the same as in Example 1 and $\alpha = \frac{2}{3}$. Then R_0 type α-reverse triple I FMP sustaining solution B^* does not exist. In fact, By $(A^*(x))' = \frac{1}{2}$ and $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$, (3) is

$$\begin{aligned} & R_0(R_0(A^*(x), B^*(y)), R_0(A(x), B(y))) \\ = & R_0(R_0(\frac{1}{2}, B^*(y)), \frac{1}{2}) \\ = & \begin{cases} \frac{1}{2}, & \frac{1}{2} \leq B^*(y) \\ 1, & \frac{1}{2} > B^*(y) \end{cases} \\ \geq & \frac{2}{3}. \end{aligned}$$

If the above inequation holds, then $B^*(y) < \frac{1}{2}$. So the maximal $B^*(y)$ does not exist. That is, there does not exist the maximal fuzzy set B^* satisfying (3).

Thus, R_0 type α-reverse triple I FMP sustaining solution does not exist.

Sustaining principle of α-reverse triple I FMT [7]. Suppose that X and Y are nonempty, $A \in \mathcal{F}(X), B, B^* \in \mathcal{F}(Y)$. Then A^* satisfying this principle is the minimal fuzzy set satisfying (3).

Below, we illustrate that R_0 type α-reverse triple I FMT sustaining solution does not exist.

Example 4 Suppose that X, Y, A, B, B^* are the same as in Example 3 and $\alpha = \frac{2}{3}$. Then for any $x \in X$ and any $y \in Y$, there does not exist the minimal fuzzy set A^* satisfying (3). In fact, (3) is represented by

$$R_0(R_0(A^*(x), \frac{1}{2}), \frac{1}{2}) = \begin{cases} \frac{1}{2}, & A^*(x) \leq \frac{1}{2} \\ 1, & A^*(x) > \frac{1}{2} \end{cases} \geq \frac{2}{3}.$$

By the above inequation, we have $A^*(x) > \frac{1}{2}$. So the minimal A^* satisfying (3) does not exist.

Hence, R_0 type α-reverse triple I FMT sustaining solution does not exist.

2.3 Existence of R_0 Type Reverse Triple I Restriction Solutions

Now we consider the existence of R_0 type reverse triple I restriction solution.

Restriction principle of reverse triple I FMP [7]. Suppose that X and Y are non-empty, $A, A^* \in \mathcal{F}(X), B \in \mathcal{F}(Y)$. Then $B^* \in \mathcal{F}(Y)$ satisfying this principle is the minimal fuzzy set that makes (2) take its minimum for any $x \in X$ and any $y \in Y$.

The following example shows that R_0 type reverse triple I FMP restriction solution does not exist.

Example 5 Suppose $X = Y = [0, 1]$, $A(x) = \frac{1}{2}$, $A^*(x) = \frac{2}{3}$, $B(y) = \frac{1}{3}$. Then for any $x \in X$ and any $y \in Y$, there does not exist the minimal fuzzy set B^* that makes (2) take its minimum. In fact, we have $(A^*(x))' = \frac{1}{3}$ and $R_0(A(x), B(y)) = R_0(\frac{1}{2}, \frac{1}{3}) = \frac{1}{2}$. By the property of R_0, the minimum of (2) is $R_0(R_0(A^*(x), 1), R_0(A(x), B(y))) = R_0(1, R_0(\frac{1}{2}, \frac{1}{3})) = \frac{1}{2}$. We analyze whether there exists the minimal fuzzy set B^* that makes (2) take its minimum. By the definition of R_0, (2) is expressed by

$$R_0(R_0(\frac{2}{3}, B^*(y)), R_0(\frac{1}{2}, \frac{1}{3})) = R_0(R_0(\frac{2}{3}, B^*(y)), \frac{1}{2}) = \begin{cases} 1, & R_0(\frac{2}{3}, B^*(y)) \le \frac{1}{2} \\ \frac{1}{2}, & R_0(\frac{2}{3}, B^*(y)) > \frac{1}{2} \end{cases}.$$

So, if (2) takes its minimum $\frac{1}{2}$, then $R_0(\frac{2}{3}, B^*(y)) > \frac{1}{2}$. That is,

$$R_0(\frac{2}{3}, B^*(y)) = \begin{cases} 1, & \frac{2}{3} \le B^*(y) \\ B^*(y), & \frac{2}{3} > B^*(y) \ge \frac{1}{3} \\ \frac{1}{3}, & \frac{1}{3} > B^*(y) \end{cases}$$
$$> \frac{1}{2}.$$

From the above inequation, we have $B^*(y) > \frac{1}{2}$. Thus, there does not exist the minimal $B^*(y)$ that makes (2) take its minimum.

By this example, we obtain that R_0 type reverse I FMP restriction solution does not always exist.

Similarly, we consider the problem of FMT.

Restriction principle of reverse triple I FMT [7]. Suppose that X and Y are non-empty, $A \in \mathcal{F}(X)$, $B, B^* \in \mathcal{F}(Y)$. Then $A^* \in \mathcal{F}(X)$ satisfying this principle is the maximal fuzzy set that makes (2) take its minimum for any $x \in X$ and any $y \in Y$.

First, we illustrate that R_0 type reverse triple I FMT restriction solution does not exist.

Example 6 Suppose $X = Y = [0, 1]$, $A(x) = \frac{2}{3}$, $B(y) = \frac{1}{2}$, $B^*(y) = \frac{1}{3}$. Then for any $x \in X$ and any $y \in Y$, there does not exist the maximal fuzzy set A^* that makes (2) take its minimum. In fact, we get $(B^*(y))' = \frac{2}{3}$ and $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$. Since R_0 is monotonically decreasing with respect to the first variable, the minimum of (2) is $R_0(R_0(0, B^*(y)), R_0(A(x), B(y))) = R_0(1, R_0(A(x), B(y))) = R_0(A(x), B(y)) = \frac{1}{2}$. And (2) is expressed by

$$R_0(R_0(A^*(x), \frac{1}{3}), \frac{1}{2}) = \begin{cases} 1, & R_0(A^*(x), \frac{1}{3}) \leq \frac{1}{2} \\ \frac{1}{2}, & R_0(A^*(x), \frac{1}{3}) > \frac{1}{2} \end{cases}.$$

Obviously, if (2) takes its the minimum $\frac{1}{2}$, then

$$R_0(A^*(x), \frac{1}{3}) > \frac{1}{2}. \tag{4}$$

By

$$R_0(A^*(x), \frac{1}{3}) = \begin{cases} 1, & A^*(x) \leq \frac{1}{3} \\ (A^*(x))', & \frac{1}{3} < A^*(x) \leq \frac{2}{3} \\ \frac{1}{3}, & A^*(x) > \frac{2}{3} \end{cases},$$

if $R_0(A^*(x), \frac{1}{3}) > \frac{1}{2}$, then $A^*(x) < \frac{1}{2}$. Obviously, there does not exist the maximal fuzzy set A^* satisfying (4). So R_0 type reverse triple I FMT restriction solution does not exist.

2.4 Existence of R_0 Type α-reverse Triple I Restriction Solutions

In this section, we consider the generalization of reverse triple I restriction methods.

Restriction principle of α-reverse triple I FMP [8]. Suppose that X and Y are non-empty, $A, A^* \in \mathcal{F}(X)$, $B \in \mathcal{F}(Y)$. For a given $\alpha \in (0, 1]$, B^* satisfying this principle is the minimal fuzzy set in $\mathcal{F}(Y)$ satisfying

$$(A^*(x) \longrightarrow B^*(y)) \longrightarrow (A(x) \longrightarrow B(y)) \leq \alpha \tag{5}$$

for any $x \in X$ and any $y \in Y$.

The following example shows that R_0 type α-reverse triple I FMP restriction solution does not exist.

Example 7 Suppose $X = Y = [0, 1]$, $A(x) = \frac{2}{3}$, $A^*(x) = \frac{2}{3}$, $B(y) = \frac{1}{2}$ and $\alpha = \frac{1}{2}$. Then there does not exist the minimal fuzzy set B^* satisfying (5). In fact, by $(A^*(x))' = \frac{1}{3}$ and $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$, if (5) holds, then

$$R_0(R_0(\frac{2}{3}, B^*(y)), \frac{1}{2}) \leq \frac{1}{2}. \tag{6}$$

As

$$R_0(R_0(\frac{2}{3}, B^*(y)), \frac{1}{2}) = \begin{cases} 1, & \frac{1}{2} \geq B^*(y) \\ \frac{1}{2}, & \frac{1}{2} < B^*(y) \end{cases},$$

if (6) holds, then $\frac{1}{2} < B^*(y)$. Obviously, the minimal B^* does not exist. Thus there does not exist the minimal fuzzy set satisfying (6).

Hence, R_0 type α-reverse triple I FMP restriction solution does not always exist.

At last, we consider the generalized problem of reverse triple I FMT restriction method.

Restriction principle of α-reverse triple I FMT [8]. Suppose that X and Y are nonempty, $A \in \mathcal{F}(X)$, $B, B^* \in \mathcal{F}(Y)$. $A^* \in \mathcal{F}(X)$ should be the maximal fuzzy set satisfying (5) for any $x \in X$ and any $y \in Y$.

Example 8 Suppose $X = Y = [0, 1]$, $A(x) = \frac{2}{3}$, $B(y) = \frac{1}{2}$, $B^*(y) = \frac{1}{3}$ and $\alpha = \frac{1}{2}$. There does not exist the maximal fuzzy set A^* satisfying (5). In fact, by $(B^*(y))' = \frac{2}{3}$ and $R_0(A(x), B(y)) = R_0(\frac{2}{3}, \frac{1}{2}) = \frac{1}{2}$, if (5) holds, then

$$R_0(R_0(A^*(x), \frac{1}{3}), R_0(\frac{2}{3}, \frac{1}{2})) \leq \frac{1}{2}. \tag{7}$$

The left side of the above inequation equals

$$R_0(R_0(A^*(x), \frac{1}{3}), \frac{1}{2}) = \begin{cases} 1, & R_0(A^*(x), \frac{1}{3}) \leq \frac{1}{2} \\ \frac{1}{2}, & R_0(A^*(x), \frac{1}{3}) > \frac{1}{2} \end{cases}.$$

Clearly, if (7) holds, then $R_0(A^*(x), \frac{1}{3}) > \frac{1}{2}$, i.e.

$$R_0(A^*(x), \frac{1}{3}) = \begin{cases} 1, & A^*(x) \leq \frac{1}{3} \\ (A^*(x))', & \frac{1}{3} < A^*(x) \leq \frac{2}{3} \\ \frac{1}{3}, & A^*(x) > \frac{2}{3} \end{cases}$$
$$> \frac{1}{2}.$$

So we get $A^*(x) < \frac{1}{2}$, and there does not exist the maximal fuzzy set A^* satisfying (5).

Thus R_0 type α-reverse triple I FMT restriction solution does not always exist.

3 Conclusion

In this paper, we illustrated that eight kinds of R_0 type reverse triple I solutions given in [7, 8] do not exist.

Acknowledgments Thanks to the support by National Natural Science Foundation of China (No. 11401030), and Teacher training Project Fund in Beijing Normal University, Zuhai.

References

1. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision process. IEEE Trans. Syst. Man Cybern. **3**, 28–44 (1973)
2. Wang, P.Z., Li, H.X.: The Theory of Fuzzy System and Fuzzy Computer. Science Press, Beijing (1995) (in Chinese)
3. Wu, W.M.: The Principle and Methods of Fuzzy Reasoning. Guizhou Science and Technology Press, Guiyang (1994) (in Chinese)
4. Zhang, W.X., Liang, G.X.: Fuzzy Control and System. Xi'an Jiaotong University Press, Xi'an (1998) (in Chinese)
5. Wang, G.J.: The triple I method with total implication for fuzzy reasoning, Sci. China Ser. E **29**(1), 43–57 (1999) (in Chinese)
6. Wang, G.J.: Non-classical Mathematical Logic and Approximate Reasoning. Science Press, Beijing (2000) (in Chinese)
7. Song, S.J., Wu, C.: Reverse triple I method for fuzzy reasoning. Sci. China Ser. E **32**(2),230–246 (2002) (in Chinese)
8. Song, S.J., Wu, C.: Reverse triple I restriction method for fuzzy reasoning. Progr. Nat. Sci. **12**(1), 95–100 (2002) (in Chinese)

Research on the Characteristic of the Five-Element Sub-lattice

Hai-feng Zhang, Meng Zhou and Guang-jun Zhang

Abstract We have already known in the sense of isomorphism, there are five five-element lattices. In this paper, we establish some important results about the characteristic of a lattice that contains these five five-element lattices.

Keywords Lattice · Five-element lattice · Distribution lattice · Modular lattice · Order sublattice

1 Introduction

Lattice algebra is a special algebraic system, it has been widely used in various fields. The researches on the lattice's own structure are also very important. So far, people have got many valuable conclusions (see [1, 3]). Especially using a lattice's own subalgebra characteristic to deduce its structure has been proved to be a very effective way. Using this method we give several results below (see [4]).

Theorem 1 *Let L be a lattice. L is distributive lattice if and only if L has no sublattice isomorphic to* $\mathscr{N}_5$ *or* $\mathscr{M}_3$.

Theorem 2 *(i) A lattice L is modular if and only if it has no sublattice isomorphic to* $\mathscr{N}_5$.
(ii) A modular lattice L is nondistributive if and only if it has a sublattice isomorphic to $\mathscr{M}_3$.

These two typical examples of nondistributive lattices are $\mathscr{N}_5$ and $\mathscr{M}_3$ whose diagrams are given in the Fig. 1.

H. Zhang (✉) · M. Zhou · G. Zhang
School of Mathematics and Systems Science, School of Instrumentation Science and Opto-electronics Engineering, Beihang University, XueYuan Road No. 37, Haidian, Beijing 100191, China
e-mail: zhfcun@163.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_58

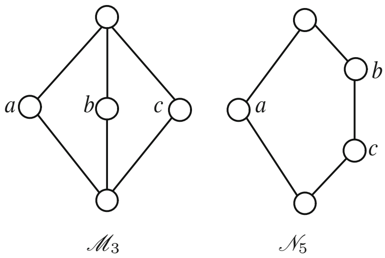


Fig. 1 Lattice $\mathscr{M}_3$, $\mathscr{N}_5$

In this paper, we will use the results above to discuss a lattice L which has a sublattice isomorphic to a specific five-element lattice, and a serial of important results can be deduced.

2 Several Definitions and Conclusions

In this section we introduce some definitions which are used in this paper.

Definition 1 (*sublattice*) A sublattice $\langle K; \wedge, \vee \rangle$ of the lattice $\langle L; \vee, \wedge \rangle$ is defined on a nonvoid (see [2]) subset K of L with the property that $a, b \in K$ implies that $a \vee b, a \wedge b \in K$, and the $\vee$ and the $\wedge$ of $\langle K; \wedge, \vee \rangle$ are restrictions to K of the $\vee$ and the $\wedge$ of $\langle L; \wedge, \vee \rangle$.

Definition 2 (*order sublattice*) Let $\langle L, \leq \rangle$ be a lattice, A is the nonvoid subset of L. If the poset$(A. \leq_A)$ is a lattice, then we call$(A. \leq_A)$ is a order sublattice of L. $\leq_A = \leq |_A$ here, that is, the $\leq_A$ of$(A. \leq_A)$ is restrictions to A of the $\leq$ of $(L, \leq)$.

Example 1 Let $L = \{0, 1, a, b, c\}$ see Fig. 2. It can be proved that $A = \{a, 0, 1, b\}$ is a order sublattice of L, but not a sublattice of L. Because $a \wedge b = c \notin A$.

The above example shows that an order sublattice of a lattice may not be a sublattice of the lattice. Let $\langle L, \leq \rangle$ be a lattice and L_1 be a sublattice of L. $\forall a, b \in L_1$,

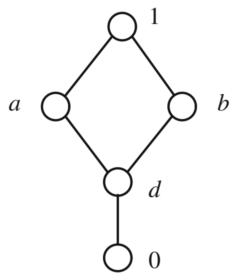


Fig. 2 Lattice $\underline{K_5}$

$a \wedge b = inf\{a, b\}$. $inf\{a, b\}$ here is the greatest lower bound of $\{a, b\}$ under the partial ordering $\leq$ of L. And $a \vee b = sup\{a, b\}$. $sup\{a, b\}$ here is the least upper bound of $\{a, b\}$ under the partial ordering $\leq$ of L.

Theorem 3 *Every sublattice of a lattice L is an order sublattice of L.*

So the problem is : under what conditions would an order sublattice of a lattice be a sublattice? This problem is solved in Theorem 3. Now we define the length of a lattice.

Definition 3 (*the length of a lattice*) Let L be a lattice. If any chain in L has size $h \leq m$ and there is a chain in L has size $h = m(0 < m < +\infty)$, then we say that L is of length $m - 1$. The length of L is denoted by $l(L)$. If the number of elements in a chain of L is infinite, we say that $l(L) = \infty$.

Obviously in Fig. 2 the length of L is $l(L) = 3$. In Fig. 1 the length of $\mathscr{N}_5$ is $l(\mathscr{N}_5) = 3$ and the length of $\mathscr{M}_3$ is $l(\mathscr{M}_3) = 2$.

Theorem 4 *In the sense of isomorphism, there are five five-element lattices.*

3 Main Results

Theorem 5 *Let* $(L. \leq)$ *be a lattice. L has a sublattice isomorphic to* C_5 *in Fig. 3 if and only if* $l(L) \geq 4$.

Theorem 6 *Let* $(L. \leq)$ *be a lattice. L has a finite chain of n elements if and only if* $l(L) \geq n - 1$.

Theorem 7 *Let* $(L. \leq)$ *be a lattice. L dose not contain* $\overline{K_5}$ *or* $\underline{K_5}$ *in Fig. 3 as a sublattice of L if and only if every order sublattice of L is a sublattice of L.*

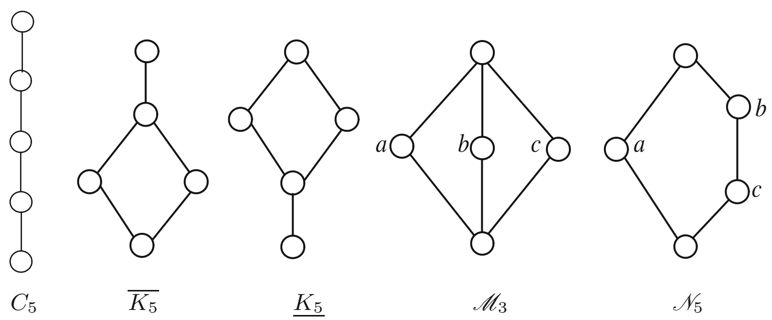


Fig. 3 Five-element lattices

Proof ($\Longrightarrow$)

Because L does not contain $\overline{K_5}$ or $\underline{K_5}$ as a sublattice of L. Suppose A is an order sublattice of L but A is not a sublattice of L, then $\exists a, b \in A$,

$$a \wedge b \notin A \text{ or } a \vee b \notin A$$

Thus $\inf_A\{a, b\} < a \wedge b$ or $\sup_A\{a, b\} > a \vee b$. And so $K_1 = \{a, b, a \vee b, a \wedge b, \inf_A\{a, b\}\}$ or $K_2 = \{a, b, a \wedge b, a \vee b, \sup_A\{a, b\}\}$ is a sublattice of L. This contradicts the assumption. Hence A is a sublattice of L.

($\Longleftarrow$)

Suppose $\underline{K_5} = \{a \vee b, a, b, a \wedge b, c\} \subseteq L$ and c is the least element of $\underline{K_5}$.

By the definition of order sublattice, $A = \{a \vee b, a, b, c\}$ is order sublattice of L. Since $c \neq a \wedge b$, we know $A = \{a \vee b, a, b, c\}$ is not a sublattice of L. This contradicts the assumption. Hence L dose not contain $\underline{K_5}$ as a sublattice of L. Similarly for $\overline{K_5}$ in Fig. 3.

In conclusion L dose not contain $\overline{K_5}$ or $\underline{K_5}$ in Fig. 3 as a sublattice of L if and only if every order sublattice of L is a sublattice of L. □

Theorem 8 *Let* $(L. \leq)$ *be a lattice. L dose not contain* $\overline{K_5}$ *or* $\underline{K_5}$ *in Fig. 3 as a sublattice of L if and only if L is a chain or L is a bounded lattice and* $\forall a, b \in L$*, if* a, b *are incomparable, then a is a complement of b (that is,* $a \wedge b = 0$ *and* $a \vee b = 1$*,* 0 *is a zero of L and* 1 *is a unit of L).*

Proof ($\Longrightarrow$)

If L is a chain. This proof is obvious directly.

If L is not a chain, then $\exists a, b \in L$ and a, b are incomparable. $\{a \wedge b, a \vee b, a, b\}$ is a four-element lattice.

Since L dose not contain $\overline{K_5}$ or $\underline{K_5}$ in Fig. 3 as a sublattice of L.

For $\forall c \in L$ we must have $a \wedge b \leq c \leq a \vee b$. So $a \wedge b = 0$ and $a \vee b = 1$.

($\Longleftarrow$)

If L is a chain, then obviously L dose not contain $\overline{K_5}$ or $\underline{K_5}$ in Fig. 3 as a sublattice of L.

If L is a bounded lattice, then there exist $0, 1 \in L$ and 0 is a zero of L, 1 is a unit of L.

Suppose L contains $\underline{K_5} = \{a, b, a \wedge b, a \vee b, c\}$. c is the least element of $\underline{K_5}$. Obviously a, b in L are incomparable, thus $a \wedge b = 0$. But $c < a \wedge b$, that is, $c < 0$. This contradicts the assumption. So L dose not contain $\underline{K_5}$ in Fig. 3 as a sublattice of L. Similarly for $\overline{K_5}$ in Fig. 3.

In conclusion L dose not contain $\overline{K_5}$ or $\underline{K_5}$ in Fig. 3 as a sublattice of L if and only if L is a chain or L is a bounded lattice and $\forall a, b \in L$, if a, b are incomparable, then a is a complement of b. □

By Theorem 8, we have some useful corollaries below.

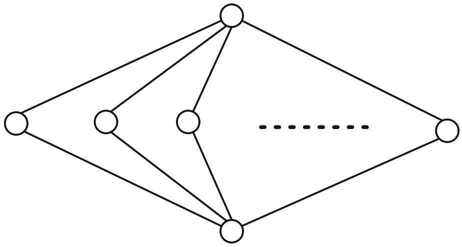


Fig. 4 Generalized diamond $\mathscr{M}$

Corollary 1 *If L is distributive lattice, then there are only two situations that* $\underline{K_5}$ *and* $\overline{K_5}$ *are not sublattice of L:*

(1) L is a chain.
(2) L is $C_2 \times C_2$*.*

Corollary 2 *If* L, K *are lattices, suppose* $|L| > 1, |K| > 1$*, then the five-element lattices* $\underline{K_5}$ *and* $\overline{K_5}$ *must be the sublattice of direct product lattice* $L \times K$*.*

Corollary 3 *If L is not distributive modular lattice, then the following conditions are equivalent:*

(1) $\underline{K_5}$ *and* $\overline{K_5}$ *are not five-element sublattice of L,*
(2) the length of L is $l(L) = 2$*,*
(3) L is a generalized diamond $\mathscr{M}$ *(see Fig. 4).*

4 Concluding Remarks

We summarize the above results and we can get a complete description on whether a specific five-element lattice C_n, $\mathscr{N}_5$, $\mathscr{M}_3$, $\underline{K_5}$ or $\overline{K_5}$ belongs to L.

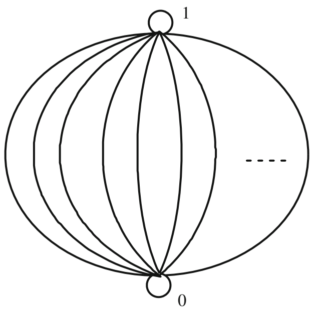


Fig. 5 A lattice does not include $\underline{K_5}$ or $\overline{K_5}$

Theorem 9 *(1)* C_n *is a sublattice of lattice L if and only if* $l(L) \geq n - 1$.
(2) $\mathscr{N}_5$ *is a sublattice of lattice L if and only if L is not modular lattice.*
(3) $\mathscr{M}_3$ *is a sublattice of lattice L if and only if L is not distributive modular lattice.*
(4) $\underline{K_5}$ *or* $\overline{K_5}$ *is a sublattice of L if and only if L is neither a chain, nor a lattice which is composed of the maximum element or the minimum element of some bounded chains (see Fig. 5).*

Acknowledgments This research was financially supported by the National Natural Science Foundation of China (No. 11271040 and 10871017).

References

1. Baker, K.: Equational axioms for classes of lattices. Bull. Am. Math. Soc. **77**, 97–102 (1971)
2. Birkhoff, G.: Lattice Theory, vol. 25, 3rd edn. American Mathematical Society Colloquium Publications, Providence (1967). MR 37 2638
3. Freese, R.: Varieties generated by modular lattices of width four. Ph.D. Thesis, California Institute of Technology (1972)
4. Gratzer, G.: General Lattice Theory. Academic Press, New York (2003)

Remote Monitoring System of River Basin Water Quality Parameters Based on LabVIEW

Qianjun Xiao, Qian Wu and Hu Xu

Abstract Aimed at the puzzle of being difficult to make the real-time online detection for water parameter field of the river basin, the paper designed a sort of remote monitoring system of environment parameter based on LabVIEW. In the paper, it explored the implementation structure of remote monitoring system under scurviness environment, discussed the advantages of virtual instrument technology, based on the combination of wireless sensor network and LabVIEW, it took the real-time online detection of water temperature parameter as an example, designed the sensing system of temperature, the architecture of the remote monitoring system and the application software system. Using the data information transmission mode of GPRS, it implemented a remote monitoring system of environmental parameters of the virtual instrument based on LabVIEW. The system response of simulation demonstrated that it would be high in its test accuracy. The research result shows that the design of the system is feasible and effective.

Keywords Virtual instrument technology · LabVIEW · Temperature sensing system · GPRS · Remote monitoring system

Q. Xiao (✉)
School of Automation, Chongqing Industry Polytechnic College,
Chongqing 401120, China
e-mail: Xiaoqianjun2003@126.com

Q. Wu
School of Computer Science and Engineering, Chongqing University
of Technology, Chongqing 400054, China
e-mail: wuqian80@163.com

H. Xu
College of Automation, Chongqing University, Chongqing 400044, China
e-mail: 1597725908@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_59

1 Introduction

The Monitoring of water quality parameters plays a significant role in the water environmental protection, and the practical significance and its importance is self-evident. The water quality parameters reflect features of the physical state of the water environmental system, it comprehensively reflects the quality of the complex ecological environment, and the most direct method of understanding the features in water quality status is to make the real-time measurement of water parameter state. The method of parameter detection based on traditional instruments is difficult to achieve the desired detection effect in most cases, and its reason maybe is that the working environment is hard to wire connection, the vast is inconvenient to install detection instrumentation, or the detect signal is difficult to transmission because of the field interference being too large and so on. For example, the monitoring of river basin water quality is difficult or not using conventional instruments to make the monitoring. With the speedy development of modern sensor, microelectronics, communication network, intelligent technology, computer information and other advanced technology, it provides a tremendous development space for that the monitoring parameter of river basin water quality can be tracked in real-time online by means of remote-networking, and has made many achievements of application [1–4], but there is still the space for further improvement. The following briefly explores a kind of remote monitoring system of water quality parameters based on LabVIEW.

2 Architecture of Remote Monitoring System Based on LabVIEW

The advantage of Virtual Instrument Technology is that it can overcome the limitations of conventional defect instrument applications [5–7]. For example, it can achieve specific detection function by means of the combination of different functional modules, and the user can defines the instrument function according to need in flexible mode. The software module concentrated by control information can enhance the flexibility of application, and it is convenient for the user to display the collected data, analyzed result and monitored process. It is conducive to data sharing because of the link among network, peripherals and other applications adopting the whole software pattern, and being small effected by hardware limitations. It can make the data transmission by using wired or wireless mode for edited data record. The powerful graphical user interface can make the analysis processing for detection data in real time. It can save the cost of development and maintenance based on software architecture. Except having the same compatibility of structure component, the virtual instrument is different from conventional instrument in the aspect of the structure framework. It has the following characteristics that it does not emphasize the realizing form of physical detection, and it

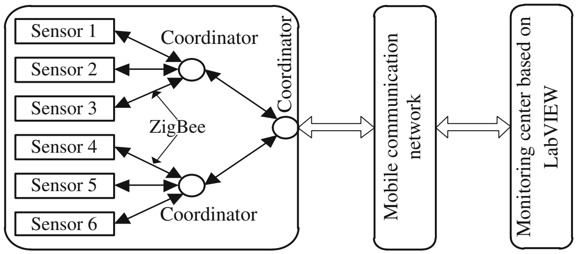


Fig. 1 Architecture of monitoring system based on LabVIEW

highlights the software realizing form. For example, it is realized by way of application software such as data acquisition, data processing, test result, data analysis and process display and so on. The detection system is based on the resource sharing of hardware and software, and under the same hardware conditions, it defines the detection function by means of software programming. Using graphical software panel as being similar to the actual instrument knob, the user can inspect the communication and operation of test instruments with the help of keyboard or mouse. The remote monitoring system based on LabVIEW is composed by the sensor, wireless network node and host computer of upper supervision [8–12], and the basic framework of detection system is shown as in Fig. 1.

The sensor is responsible for detecting physical quantity being transformed into electrical signals, and the signal is transmitted to the ZigBee coordinator with the help of wireless MESH network. Then through the serial port via the mobile communication network, the electrical parameter signal is sent to the monitoring center based on LabVIEW. The architecture is easy to realize the transmission of field detection parameter by the wireless mobile communication network with GPRS mode. The monitoring center can conveniently realize the wide range of distributed monitoring. The architecture strips the function of conventional instrument, but it also contains the function of conventional measuring instruments such as data acquisition, data processing, data analysis and data display. Because the structure of the underlying hardware platform has provided the application program interface, but it makes no limit for the upper software of monitoring center, and therefore, it can conveniently make the design and deployment for various applications with the help of LabVIEW graphical programming software design system.

3 Design of Remote Monitoring System

3.1 System of Parameters Sensing Transformation

The design principle of parameter sensing transformation system is the same, and the difference is only that the sensor is different. In this paper, it only takes the transformation of the temperature sensing system as an example, and briefly describes its transformation principle. For example, based on the linear relationship

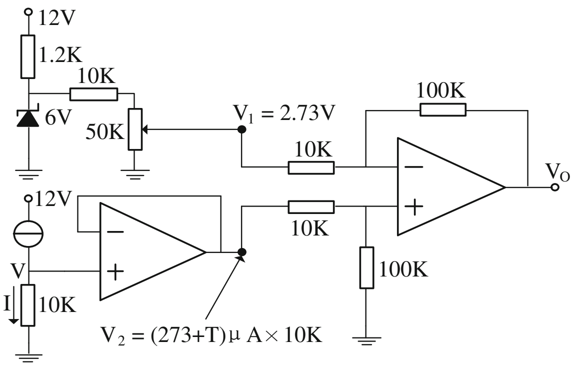


Fig. 2 Signal conditioning circuit

between the forward current and temperature in PN junction, it selects the temperature sensor of current output type AD590, and if the temperature increases 1 °C, then the current increases 1 μA in the range from −55 °C to 150 °C. The signal conditioning circuit is shown as in Fig. 2.

The current produced by the device is proportional to absolute temperature, and it is equivalent to a constant current source. The conditioning circuit makes the temperature map as its corresponding current value, eliminates the influence of the power supply variations and it is directly converted to voltage value. From the Fig. 3, it can be seen that the output current I = (273 + T) μA (T is the Celsius temperature), and therefore, the measurement voltage is V = (273 + T) μA × 10 kΩ = (2.73 + T/100) V. The output circuit adopts the voltage follower of high input impedance, and so the output voltage is V2 = V. In order to simplify the computation, the voltage V1 will be adjusted to 2.73 V, and then the output of differential amplifier is (100 kΩ/10 kΩ) × (V2 − V1) = T/10 V,such as T = 28, Vo = 2.8 V.

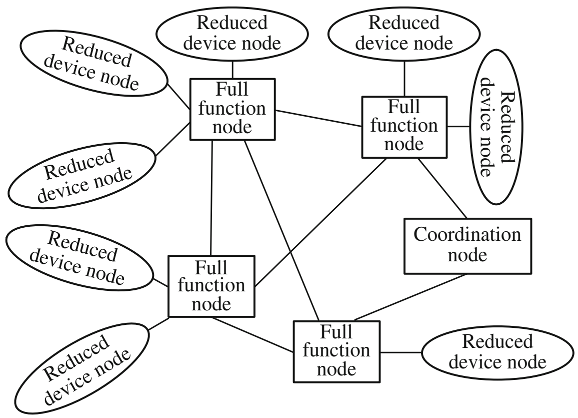


Fig. 3 Topology structure of mesh network

3.2 Design of Wireless Sensor Network

The sketch map of mesh topology of wireless sensor network is shown as in Fig. 3. In Fig. 3, there are three types of nodes. For full function node, it is mainly responsible for the collection of physical parameters such as temperature etc. and its physical address of the node as well as receiving the command sent by coordination node. But it has the routing relay function, namely it is not only the monitor terminal, but also it owns the full function device node with wireless transmission device. The reduced device node is the terminal device for simplified detecting node. The difference compared with the full function node is that it is the device terminal node without routing relay function. The network coordinator is the coordination node of wireless sensor networks, and it is responsible for receiving and processing the data for each acquisition node. Through the RS232 serial port, it makes the data exchange with upper host computer, and it is responsible for starting and building the network. In the scope of coverage in monitoring network, a coordinating node (access point) must be set up, and with the help of the node, it can access any node in the mesh network. The type and number of wireless sensor network node configuration are determined by monitoring regional range. As long as there is a communication signal, and each full function node can transmit the data by means of relay node. With the help of coordinator node, it can ensure the connection with the mobile communication network. Once a certain detection node appears the failure, the neighboring node can send the test data to the network coordinator through the relay node, and by means of GPRS, the data information would be transmitted to the monitoring center.

3.3 Communication Mode

General Packet Radio Service (GPRS) is a sort of wireless block-switch technology based on Global System for Mobile Communications (GSM), and it provides the wireless communication connection of peer-to-peer. By means of communication service of GPRS, it can make the data of network Coordinator realize the communication with upper layer software of monitoring center using the transmitting and receiving of sub-group mode. Adopting this communication mode, it is more convenient to make the communication connection and to send the data information because of being almost no need of time to build the new connection. When the data is sent, the communication channel is still on line all the time, and it can reduce the cost in large range. The communication of GPRS charges according to the data throughput, and the paid communication cost is just the transfer fee of data packet, but not the cost of the entire communication channel. In addition, the GPRS data transfer rate is high, and the data transfer rate of GPRS is 10 times data transfer rate of GSM, it can reach to 171.2 kbps, and therefore, it is desirable to select the data transfer mode of GPRS.

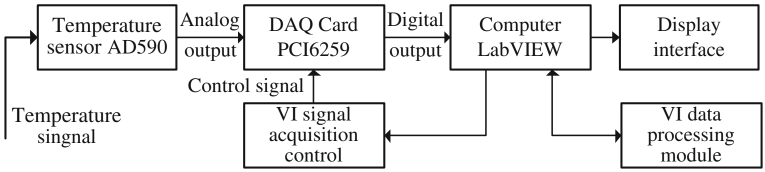


Fig. 4 Description of LabVIEW based detection principle

3.4 LabVIEW Based Monitoring Center

The LabVIEW is a sort of advanced system design software, and it adopts the intuitive graphical method to design the system. By means of graphical programming environment, it can conveniently to carry out the design and deployment of various applications. The principle of various physical parameters detection is similar, and the different sensor can form different virtual instrument VI module. Here, taking only the water temperature parameter detection as an example, the detection process of the whole parameters is briefly described. The temperature detection process is composed of the temperature sensor, instrument hardware and computer and so on, and the schematic diagram is shown as in Fig. 4. The output of temperature sensor AD590 is sent to the monitoring host computer through wireless sensor network by means of GPRS, the LabVIEW uses complete software mode to carry out the processing for the acquisition data, and finally, it makes the detection result be displayed directly in the virtual instrument interface of the monitoring center. With the help of monitoring interface and data browsing interface, it can understand the change situation of dynamic distribution in time for various physical parameters in monitored region scope, and show the environmental state by the mode of visual dynamic graphics in intuitive image and real time curve graph and so on. Of course, it can also query and analyze historical data and so on.

4 Engineering Implementation of Remote Monitoring System

The main parameter reflected the situation of water quality in river basin is the parameters such as pH value, temperature, dissolved oxygen and heavy metal content and so on. After building the remote monitoring system architecture and designing the system, its engineering implementation is relatively easy [13–15]. In implementation, it builds the hardware system according to the system design, after the completing the software programming and debugging, and it can build a remote monitoring system of environmental parameters based on LabVIEW. The system can realize the remote monitoring of water environment parameters, and it is an unattended detection system that can automatically find the abnormal data and track

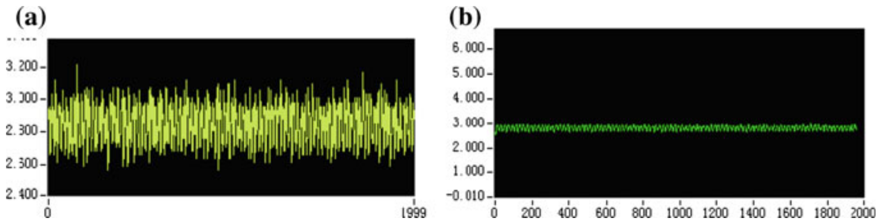


Fig. 5 Comparison of results *before* and *after treatment*. **a** Actual acquisition signal. **b** Optimized signal

Fig. 6 Real-time temperature change

the parameter change. It can also make the real-time detection for industrial wastewater emissions, and provide the first-hand raw data for the water quality management department. The Fig. 5 gives a water temperature signal of field monitoring temperature sensor AD590, and in which, the Fig. 5a is the signal of actual temperature sensor, while the Fig. 5b is the optimized signal after denoising processing by the LabVIEW. Through the contrast of Fig. 5a, b, it can be seen that the LabVIEW has powerful signal processing functions. The Fig. 6 is the real-time temperature curve of a certain monitoring point in a time period of some day. It is the same to the monitoring situation of other parameters such as pH value, dissolved oxygen and heavy metal content and so on, and here it is no longer explained.

The experimental test verifies the feasibility and rationality of environmental parameters remote monitoring system based on LabVIEW.

5 Conclusions

Based on the combination of respective technological advantage of LabVIEW and GPRS, the above explored the related design of the remote monitoring system in water quality parameters. The proposed scheme is good for saving instrument hardware device and system running communication cost. It can suppress the acquisition noise by software mode, and it owns the flexibility for system function module design. It is accurate and reliable in measurement, rapid in response, powerful in function and better in extendibility, and therefore, it is a relatively ideal choice of options for monitoring of water quality in the river basin.

References

1. Zhang, H.T., Nie, S.L.: Design of remote monitoring system for environmental parameters based on GPRS technology. Microcomput. Appl. **29**(22), 96–99 (2010)
2. Wang, H.B., Wu, T.T., Wu, G.J.: Air quality monitoring system based on LabVIEW. Comput. Meas. Control **19**(3), 525–527 (2011)
3. Xie, W.J., Ding, Y.M., Dong, J., et al.: On-line monitoring system with remote access function for grouping welding processes based on LabVIEW. Heat Treat. Technol. **41**(01), 147–150 (2012)
4. Wang, X.H., Su, H.S.: Remote power monitoring system based on GPRS. J. Lanzhou Jiaotong Univ. **27**(4), 138–140 (2008)
5. Wang, C.J., Chen, X.X.: Design of the signal detection system based on the LabVIEW. Meas. Control Technol. **30**(3), 18–21 (2011)
6. Wang, S.F., Du, F.M.: Design of multi-channel data acquisition and processing system based on virtual instrument. Mach. Tool Hydraulics **37**(18), 306–308 (2009)
7. Zhao, Q.Y.: Multi-functional virtual instrument system based on digital signal processing. Industry and Mine Automation **2010**(4), 130–132 (2010)
8. Dai, X.: Study on environmental parameters remote monitoring system. Hubei Forest Ind. Technol. **2013**(2), 29–31 (2013)
9. Sun, L.M., Li, J.Z., Chen, Y.: Wireless Sensor Networks. Tsinghua University Press, Beijing (2005)
10. Li, Fan, Yang, W.J., Fang, Q.: Design of wireless sensor network node for measurement shock waves field. J. Sichuan ordnanc **35**(6), 94–97 (2014)
11. Gao, Y.Z., Xue, X.L., Gao, L.: Analysis of tracking ability of the target on the boundary of wireless sensor networks. J. Sichuan ordnanc **35**(2), 112–113 (2014)
12. Wang, S.: Theory and application of wireless sensor network. Beihang Univ. Press, Beijing (2007)
13. Liu, Y., He, K., Yuan, W.Q., et al.: Design of monitoring system for air compressor based on LabVIEW. Coal Mine Mach. **35**(12), 35–37 (2014)
14. Lv, T.G., Liu, Z.Z., Wang, S.S., et al.: The remote monitoring system for greenhouse parameters based on LabVIEW. Chin. Agric. Mechanization **2010**(2), 80–82 (2010)
15. Dai, G., Liu, P., Zhang, Y.: Remote monitoring system for pipe leakage based on LabVIEW shared variable. Instrumental Tech. **2009**(1), 15–17 (2009)

Quantum Dialogue for Networks with Partially Entangled States

Yi Xiang, Zhi-wen Mo and Shu-qin Sun

Abstract In this paper, we proposed a quantum dialogue scheme for satellite-to-ground networks with partially entangled states. Any two communication clients to establish a quantum channel via entanglement swapping with partially entangled states, which is performed by the satellite ground station. Through using Bell-basis measurement, adding auxiliary, implementing specially defined unitary matrices and appropriate Pauli matrices, the two communication clients can exchange their target qubit simultaneously with a certain probability.

Keywords Quantum dialogue · Satellite-to-ground networks · Partially entangled states · Entanglement swapping · Bell-basis measurement

1 Introduction

Quantum communication [1] is a new research field developed rapidly in recent years, which is based on the properties of quantum mechanics, rather than the high complexity of the underlying mathematical problems. In 2004, Nguyen [2] proposed the first quantum dialogue (QD) protocol, which permits two parties transmit their different secret messages to each other only in one quantum channel. After that, Some quantum dialogue schemes based on the maximally entangled states, such as Bell states, GHZ states, etc., have been extensively studied [3, 4]. However, due to the inevitable interaction of the qubits with the environment, direct transmission is not practicable for long-distance quantum communication. So it is necessary to establish quantum communication over wide area networks.

Y. Xiang (✉) · Z. Mo
College of Mathematics and Software Science, Sichuan Normal University,
Chengdu 610066, China
e-mail: xiangyiaoxiang@163.com

S. Sun
Department of Mathematics, Sichuan Minzu College,
Kangding 626001, China

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_60

In recent years, many efforts have been devoted to the search of quantum information processing with partially entangled states [5, 6]. However, little has been done in the study of transferring qubits in quantum networks with partially entangled states. Moreover, existing research [7–10] on complicated quantum communication networks are also based on the maximally entangled states, ignoring the impact of decoherence effects on quantum states. Recently, Chen et al. [11] proposed a quantum communication for satellite-to-ground networks with partially entangled states. In their scheme, a quantum channel can only be used for one-way transmission of quantum bits. However, we mainly consider how to implement two-way transmission of quantum bits by one quantum channel in wide area networks.

The rest of this paper is organized as follows. In Sect. 2, we introduced two forms of entanglement swapping with partially entangled states and the satellite-to-ground quantum communication network model are briefly described. In Sect. 3, Two kinds of quantum dialogue utilized partially entangled states in satellite-to-ground quantum networks are investigated in detail.

2 Introduction of Entanglement Swapping and Network Model

2.1 Entanglement Swapping with Partially Entangled States

As we all know, the Bell states that are used in our quantum dialogue are generally defined as

$$
\begin{aligned}
|\phi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \\
|\psi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)
\end{aligned}
\tag{1}
$$

However, due to the decoherence effects of the quantum channel, the maximally entangled state $|\phi^{+}\rangle$ will be degrade to

$$
|\Upsilon^{+}\rangle = sin\theta|00\rangle + cos\theta|11\rangle, \tag{2}
$$

where $0 \le \theta \le \frac{\pi}{2}$. When $\theta = \frac{\pi}{4}$, it is a maximally entangled state. While for $\theta = 0$ or $\frac{\pi}{2}$, there is no entanglement.

Next, we analyze two forms of entanglement swapping with partially entangled states, which are described by

$$
\begin{aligned}
&|\Upsilon^{+}\rangle_{12} \otimes |\Upsilon^{+}\rangle_{34} \\
=&\frac{\sqrt{sin^4\theta + cos^4\theta}}{\sqrt{2}}(|\phi^{+}\rangle_{23} \otimes |\phi_{\theta}^{+}\rangle_{14} + |\phi^{-}\rangle_{23} \otimes |\phi_{\theta}^{-}\rangle_{14}) \\
&+ sin\theta cos\theta(|\psi^{+}\rangle_{23} \otimes |\psi^{+}\rangle_{14} + |\psi^{-}\rangle_{23} \otimes |\psi^{-}\rangle_{14})
\end{aligned}
\tag{3}
$$

where

$$|\phi_\theta^\pm\rangle_{14} = \frac{sin^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|00\rangle_{14} \pm \frac{cos^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|11\rangle_{14}, \tag{4}$$

From Eq. (3), we can see that when particles 2 and 3 are projected onto the Bell states $|\psi^\pm\rangle_{23}$, the particle pair 1, 4 is projected accordingly onto the Bell states $|\psi^\pm\rangle_{14}$ with probability $(sin\theta cos\theta)^2$. Otherwise, when particles 2 and 3 are projected onto the Bell states $|\phi^\pm\rangle$, the particle pair 1, 4 is projected as well onto the partially entangled states $|\phi_\theta^\pm\rangle$ with probability $\frac{sin^4\theta+cos^4\theta}{2}$.

Another entanglement swapping is described as follows

$$\begin{aligned}
&|\phi_\theta^+\rangle_{56} \otimes |\Upsilon^+\rangle_{78} \\
=&\frac{\sqrt{sin^6\theta + cos^6\theta}}{\sqrt{2(sin^4\theta + cos^4\theta)}}(|\phi^+\rangle_{67} \otimes |\phi_{\theta^*}^+\rangle_{58} + |\phi^-\rangle_{67} \otimes |\phi_{\theta^*}^-\rangle_{58}) \\
&+ \frac{sin\theta cos\theta\sqrt{sin^2\theta + cos^2\theta}}{\sqrt{2(sin^4\theta + cos^4\theta)}}(|\psi^+\rangle_{67} \otimes |\psi_{\theta'}^+\rangle_{58} + |\psi^-\rangle_{67} \otimes |\psi_{\theta'}^-\rangle_{58})
\end{aligned} \tag{5}$$

where

$$|\phi_{\theta^*}^\pm\rangle_{58} = \frac{sin^3\theta}{\sqrt{sin^6\theta + cos^6\theta}}|00\rangle_{58} \pm \frac{cos^3\theta}{\sqrt{sin^6\theta + cos^6\theta}}|11\rangle_{58}, \tag{6}$$

$$|\psi_{\theta'}^\pm\rangle_{58} = \frac{sin\theta}{\sqrt{sin^2\theta + cos^2\theta}}|01\rangle_{58} \pm \frac{cos\theta}{\sqrt{sin^2\theta + cos^2\theta}}|10\rangle_{58}, \tag{7}$$

From Eq. (5), we can see that when particles 6 and 7 are projected onto the Bell states $|\psi^\pm\rangle_{67}$, the particle pair 5, 8 is projected accordingly onto partially entangled states $|\psi_{\theta'}^\pm\rangle_{58}$ with probability $\frac{sin^2\theta cos^2\theta(sin^2\theta+cos^2\theta)}{2(sin^4\theta+cos^4\theta)}$. Otherwise, when particles 6 and 7 are projected onto the Bell states $|\phi^\pm\rangle$, the particle pair 5, 8 is projected as well onto partially entangled states $|\phi_{\theta^*}^\pm\rangle$ with probability $\frac{sin^6\theta+cos^6\theta}{2(sin^4\theta+cos^4\theta)}$.

When $\theta = \frac{\pi}{4}$, the above discussion corresponds to entanglement swapping with maximally entangled states.

2.2 Satellite-to-Ground Network Model

As shown in Fig. 1, we introduce a satellite-to-ground networks model, which is composed of a global position database (GPD), several satellites, several local area networks and quantum communication clients. The function of each device is no longer to do detailed introduction, which has been described in Ref. [10].

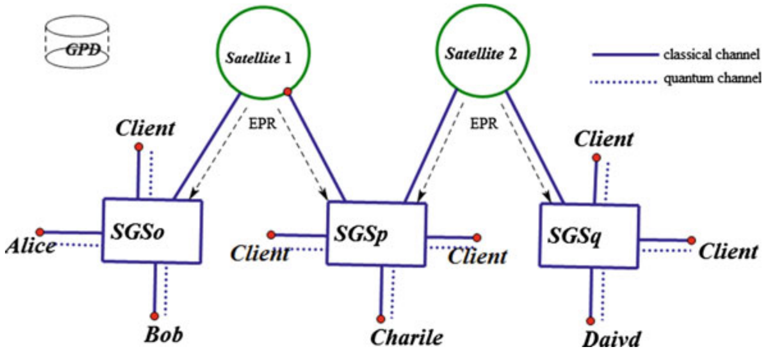


Fig. 1 Structure of satellite-to-ground quantum communication network

3 Quantum Dialogue Scheme

Before discussing the specific quantum dialogue situations, we need to make the following assumptions. we assume that each quantum dialogue client prior to share enough partially entangled states, which are uniformly described as $|\Upsilon^+\rangle$, with the SGS that serves it.

3.1 Scenario 1

The dialogue established between Alice and Bob both nodes, as shown in Fig. 1.

Firstly, Alice sends a call request with Bob to SGS_O. Then, SGS_O would query LPD_O to determine Bob's location. After applying an appropriate quantum routing protocol, the selected quantum path is Alice $\rightarrow SGS_O \rightarrow$ Bob. As the unique intermediate node, SGS_O will implement the following two entanglement swapping with $|\Upsilon^+\rangle_{A_1O_1} \otimes |\Upsilon^+\rangle_{O_3B_1}$ and $|\Upsilon^+\rangle_{A_2O_2} \otimes |\Upsilon^+\rangle_{O_4B_2}$, similarly described by Eq. (3). Here particles O_1, O_2, O_3 and O_4 belong to SGS_O, particles A_1, A_2 belong to Alice, and B_1, B_2 corresponds to Bob's. After entanglement swapping, SGS_O will send the measurement results represented by 4 bits of classical information to Bob. Then, according to the measurement results received, Bob performs a proper unitary operation from $\{I, \sigma_x, \sigma_z, i\sigma_y\}$ on particles B_1, B_2, where $I = |0\rangle\langle 0| + |1\rangle\langle 1|$, $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$, $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$, $i\sigma_y = |0\rangle\langle 1| - |1\rangle\langle 0|$. The correspondence between entanglement swapping measurement results received and local unitary operations has been shown in Table 1.

For example, if SGS_O's measurement result is $|\psi^+\rangle_{O_1O_3}(|\psi^+\rangle_{O_2O_4})$, then Bob will perform σ_x on particle $B_1(B_2)$, then the entangled state shared between himself and Alice will be $|\phi^+\rangle_{A_1B_1}(|\phi^+\rangle_{A_2B_2})$. According to Eq. (3) and Table 1, we can see that after Bob's unitary operation, the end-to-end entangled state shared between Alice and Bob will be $|\phi^+\rangle_{A_1B_1}(|\phi^+\rangle_{A_2B_2})$ or $|\phi_\theta^+\rangle_{A_1B_1}(|\phi_\theta^+\rangle_{A_2B_2})$. When $\theta = \frac{\pi}{4}$, $|\phi_\theta^+\rangle$ that is

Table 1 Relation between entanglement swapping measurement results received and local unitary operations

Measurement results	Unitary operations
$\|\phi^+\rangle$	I
$\|\phi^-\rangle$	σ_z
$\|\psi^+\rangle$	σ_x
$\|\psi^-\rangle$	$i\sigma_y$

$|\phi^+\rangle$, So we only consider the end-to-end entangled state shared between Alice and Bob is $|\phi_\theta^+\rangle_{A_1B_1}$ and $|\phi_\theta^+\rangle_{A_2B_2}$.

Now, we assume that Alice would like transfers the qubit is $|X\rangle = (a|0\rangle + b|1\rangle)_X$, Bob would like transfers the qubit is $|Y\rangle = (c|0\rangle + d|1\rangle)_Y$. The state of the entire system before quantum dialogue is represented by $|\xi\rangle_{A_1B_1A_2B_2XY}$, which can be written as

$$\begin{aligned}|\xi\rangle_{A_1B_1A_2B_2XY} &= |\phi_\theta^+\rangle_{A_1B_1} \otimes |\phi_\theta^+\rangle_{A_2B_2} \otimes |X\rangle \otimes |Y\rangle \\ &= (\frac{sin^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|00\rangle + \frac{cos^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|11\rangle)_{A_1B_1} \\ &\otimes (\frac{sin^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|00\rangle + \frac{cos^2\theta}{\sqrt{sin^4\theta + cos^4\theta}}|11\rangle)_{A_2B_2} \\ &\otimes (a|0\rangle + b|1\rangle)_X \otimes (c|0\rangle + d|1\rangle)_Y\end{aligned} \tag{8}$$

After Bob's unitary operation is completed, Alice performs Bell-basis measurements on particles X and A_1, Bob performs Bell-basis measurements on particles Y and B_2. The system composed of particles A_2 and B_1 will collapse into one of the following 16 kinds of results,

$$\begin{aligned}\langle\phi^\pm|_{B_2Y}\langle\phi^+|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\alpha}{k}|\Omega^+\rangle_{B_1} \otimes |\Gamma^\pm\rangle_{A_2} \\ \langle\psi^\pm|_{B_2Y}\langle\phi^+|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\beta}{k}|\Omega^+\rangle_{B_1} \otimes |\Gamma_*^\pm\rangle_{A_2} \\ \langle\phi^\pm|_{B_2Y}\langle\phi^-|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\alpha}{k}|\Omega^-\rangle_{B_1} \otimes |\Gamma^\pm\rangle_{A_2} \\ \langle\psi^\pm|_{B_2Y}\langle\phi^-|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\beta}{k}|\Omega^-\rangle_{B_1} \otimes |\Gamma_*^\pm\rangle_{A_2} \\ \langle\phi^\pm|_{B_2Y}\langle\psi^+|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\gamma}{k}|\Omega_*^+\rangle_{B_1} \otimes |\Gamma^\pm\rangle_{A_2} \\ \langle\psi^\pm|_{B_2Y}\langle\psi^+|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\delta}{k}|\Omega_*^+\rangle_{B_1} \otimes |\Gamma_*^\pm\rangle_{A_2} \\ \langle\phi^\pm|_{B_2Y}\langle\psi^-|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\gamma}{k}|\Omega_*^-\rangle_{B_1} \otimes |\Gamma^\pm\rangle_{A_2} \\ \langle\psi^\pm|_{B_2Y}\langle\psi^-|_{A_1X}\xi\rangle_{A_1B_1A_2B_2XY} &= \frac{\delta}{k}|\Omega_*^-\rangle_{B_1} \otimes |\Gamma_*^\pm\rangle_{A_2}\end{aligned} \tag{9}$$

where,

$$
\begin{aligned}
\alpha &= \sqrt{(|a|^2 sin^4\theta + |b|^2 cos^4\theta)(|c|^2 sin^4\theta + |d|^2 cos^4\theta)} \\
\beta &= \sqrt{(|a|^2 sin^4\theta + |b|^2 cos^4\theta)(|d|^2 sin^4\theta + |c|^2 cos^4\theta)} \\
\gamma &= \sqrt{(|b|^2 sin^4\theta + |a|^2 cos^4\theta)(|c|^2 sin^4\theta + |d|^2 cos^4\theta)} \\
\delta &= \sqrt{(|b|^2 sin^4\theta + |a|^2 cos^4\theta)(|d|^2 sin^4\theta + |c|^2 cos^4\theta)} \\
k &= 2(sin^4\theta + cos^4\theta)
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
|\Omega^{\pm}\rangle_{B_1} &= \frac{a sin^2\theta|0\rangle_{B_1} \pm b cos^2\theta|1\rangle_{B_1}}{\sqrt{|a|^2 sin^4\theta + |b|^2 cos^4\theta}} \\
|\Omega^{\pm}_{*}\rangle_{B_1} &= \frac{b sin^2\theta|0\rangle_{B_1} \pm a cos^2\theta|1\rangle_{B_1}}{\sqrt{|b|^2 sin^4\theta + |a|^2 cos^4\theta}} \\
|\Gamma^{\pm}\rangle_{A_2} &= \frac{c sin^2\theta|0\rangle_{A_2} \pm d cos^2\theta|1\rangle_{A_2}}{\sqrt{|c|^2 sin^4\theta + |d|^2 cos^4\theta}} \\
|\Gamma^{\pm}_{*}\rangle_{A_2} &= \frac{d sin^2\theta|0\rangle_{A_2} \pm c cos^2\theta|1\rangle_{A_2}}{\sqrt{|d|^2 sin^4\theta + |c|^2 cos^4\theta}}
\end{aligned} \tag{11}
$$

Then, Alice and Bob inform their Bell-basis measurement results each other (4 bits of classical information). In order to recover the target qubit $|X\rangle$ and $|Y\rangle$, they should prepare an auxiliary qubit $|0\rangle_{aux}$, and implement unitary matrix

$$
U_1 = \frac{1}{sin^2\theta}\begin{pmatrix} cos^2\theta & \sqrt{sin^4\theta - cos^4\theta} & 0 & 0 \\ 0 & 0 & 0 & sin^2\theta \\ 0 & 0 & sin^2\theta & 0 \\ \sqrt{sin^4\theta - cos^4\theta} & -cos^2\theta & 0 & 0 \end{pmatrix} \tag{12}
$$

on particles $A_2(B_1)$ and $|0\rangle_{aux}$. Then, Alice and Bob measures their auxiliary qubit with basis $\{|0\rangle, |1\rangle\}$, respectively. If the measurement results they get are $|0\rangle_{aux}$, Alice and Bob can recover qubit $|Y\rangle$ and $|X\rangle$ by performing a proper unitary operation on particles A_2 and B_1, respectively (Table 2). Otherwise, If they get $|1\rangle_{aux}$, no information about $|X\rangle$ or $|Y\rangle$ is obtained and transmission fails.

Table 2 Relation between other side's measurement results received and local recovery unitary operator

Other side's measurement results	Recovery unitary operator
$\|\phi^+\rangle$	U_1, I
$\|\phi^-\rangle$	U_1, σ_z
$\|\psi^+\rangle$	U_1, σ_x
$\|\psi^-\rangle$	$U_1, i\sigma_y$

For example, if Alice's measurement result is $|\phi^-\rangle_{A_1X}$ and Bob's measurement result is $|\psi^+\rangle_{B_2Y}$, according Eq. (9), Alice and Bob will implement unitary matrix U_1 on $|\Gamma_*^+\rangle_{A_2}|0\rangle_{aux}$ and $|\Omega^-\rangle_{B_1}|0\rangle_{aux}$, respectively.

$$
\begin{aligned}
U_1|\Gamma_*^+\rangle_{A_2}|0\rangle_{aux} &= U_1 \frac{dsin^2\theta|00\rangle_{A_2aux} + ccos^2\theta|10\rangle_{A_2aux}}{\sqrt{|d|^2sin^4\theta + |c|^2cos^4\theta}} \\
&= \frac{cos^2\theta}{\sqrt{|d|^2sin^4\theta + |c|^2cos^4\theta}}(d|0\rangle + c|1\rangle)_{A_2}|0\rangle_{aux} \\
&\quad + \frac{\sqrt{sin^4\theta - cos^4\theta}}{\sqrt{|d|^2sin^4\theta + |c|^2cos^4\theta}} d|1\rangle_{B_1}|1\rangle_{aux}
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
U_1|\Omega^-\rangle_{B_1}|0\rangle_{aux} &= U_1 \frac{asin^2\theta|00\rangle_{B_1aux} - bcos^2\theta|10\rangle_{B_1aux}}{\sqrt{|a|^2sin^4\theta + |b|^2cos^4\theta}} \\
&= \frac{cos^2\theta}{\sqrt{|a|^2sin^4\theta + |b|^2cos^4\theta}}(a|0\rangle - b|1\rangle)_{B_1}|0\rangle_{aux} \\
&\quad + \frac{\sqrt{sin^4\theta - cos^4\theta}}{\sqrt{|a|^2sin^4\theta + |b|^2cos^4\theta}} a|1\rangle_{B_1}|1\rangle_{aux}
\end{aligned} \tag{14}
$$

Next, Alice and Bob measures their auxiliary qubit with basis $\{|0\rangle, |1\rangle\}$, respectively. From Eq. (13), it can be seen that if Alice obtains $|0\rangle_{aux}$, then the particle A_2 collapses to $(d|0\rangle + c|1\rangle)_{A_2}$ and she only performs σ_x on particle A_2 to recover the qubit $|Y\rangle$. Otherwise, if she gets $|1\rangle_{aux}$, no information about $|Y\rangle$ is obtained. Similarly, from Eq. (14), it can be seen that if Bob obtains $|0\rangle_{aux}$, then the particle B_1 collapses to $(a|0\rangle - b|1\rangle)_{B_1}$ and he only performs σ_z on particle B_1 to recover the qubit $|X\rangle$. Otherwise, if he gets $|1\rangle_{aux}$, no information about $|X\rangle$ is still obtained. In Table 3, we have listed all the possible Bell-basis measurement results and the corresponding probability of successful transmission, so the probability of success is $P_t = 16 \times \frac{cos^8\theta}{|k|^2} = \frac{4cos^8\theta}{(sin^4\theta + cos^4\theta)^2}$, when $\theta = \frac{\pi}{4}$, $P_t = 1$.

Table 3 The possible Bell-basis measurement results and corresponding success probability

Measurement results	The probability of getting $\|0\rangle_{aux}$	Success probability
$\|\phi^\pm\rangle_{A_1X}, \|\phi^\pm\rangle_{B_2Y}$	$\frac{cos^8\theta}{\|\alpha\|^2}$	$\frac{cos^8\theta}{\|k\|^2}$
$\|\phi^\pm\rangle_{A_1X}, \|\psi^\pm\rangle_{B_2Y}$	$\frac{cos^8\theta}{\|\beta\|^2}$	$\frac{cos^8\theta}{\|k\|^2}$
$\|\psi^\pm\rangle_{A_1X}, \|\phi^\pm\rangle_{B_2Y}$	$\frac{cos^8\theta}{\|\gamma\|^2)}$	$\frac{cos^8\theta}{\|k\|^2}$
$\|\psi^\pm\rangle_{A_1X}, \|\psi^\pm\rangle_{B_2Y}$	$\frac{cos^8\theta}{\|\delta\|^2}$	$\frac{cos^8\theta}{\|k\|^2}$

3.2 Scenario 2

The dialogue established between Alice and Charlie both nodes, as shown in Fig. 1.

Similarly, Alice sends a call request with Charlie to SGS_O. Then, SGS_O would query GPD to determine Charlie is inside area P. After applying a quantum routing protocol, each involved node will be informed that the selected quantum path is Alice → SGS_O → SGS_P → Charlie. According to the description in Sect. 2, SGS_O and SGS_P can share two partially entangled states $|\Upsilon^+\rangle_{O_5P_1}$ and $|\Upsilon^+\rangle_{O_6P_2}$, which are distributed by the satellite 1. Here particles O_5 and O_6 belong to SGS_O, particles P_1 and P_2 belong to SGS_P.

After SGS_O's entanglement swapping with $|\Upsilon^+\rangle_{A_1O_1} \otimes |\Upsilon^+\rangle_{O_5P_1}$, $|\Upsilon^+\rangle_{A_2O_2} \otimes |\Upsilon^+\rangle_{O_6P_2}$, as similarly expressed by Eq. (3.3), and SGS_P's unitary operations according to Table 1, the possible entangled states shared between Alice and SGS_P is $|\phi^+\rangle_{A_1P_1}(|\phi^+\rangle_{A_2P_2})$ or $|\phi^+_\theta\rangle_{A_1P_1}(|\phi^+_\theta\rangle_{A_2P_2})$. Without loss of generality, we only consider the end-to-end entangled state shared between Alice and SGS_P is $|\phi^+_\theta\rangle_{A_1P_1}$ and $|\phi^+_\theta\rangle_{A_2P_2}$.

Then, SGS_P implements entanglement swapping with $|\phi^+_\theta\rangle_{A_1P_1} \otimes |\Upsilon^+\rangle_{P_3C_1}$, $|\phi^+_\theta\rangle_{A_2P_2} \otimes |\Upsilon^+\rangle_{P_4C_2}$, respectively, similarly described by Eq. (5). Here particles P_1, P_2, P_3 and P_4 belong to SGS_P, particles A_1, A_2 belong Alice, and C_1, C_2 correspond to Charlie's. And Charlie performs a proper unitary operation on particles C_1 and C_2 according Table 1. the possible end-to-end entangled states shared between Alice and Charlie is $|\phi^+_{\theta^*}\rangle_{A_1C_1}(|\phi^+_{\theta^*}\rangle_{A_2C_2})$ or $|\phi^+_{\theta'}\rangle_{A_1C_1}(|\phi^+_{\theta'}\rangle_{A_2C_2})$, where,

$$|\phi^+_{\theta'}\rangle = \frac{sin\theta}{\sqrt{sin^2\theta + cos^2\theta}}|00\rangle + \frac{cos\theta}{\sqrt{sin^2\theta + cos^2\theta}}|11\rangle, \tag{15}$$

So the possible end-to-end entangled states shared between Alice and Charlie has three cases which are given by the first column in Table 4. Suppose Alice would like transfers the target qubit is $|X\rangle = (a|0\rangle + b|1\rangle)_X$, Charlie would like transfers the qubit is $|Z\rangle = (e|0\rangle + f|1\rangle)_Z$. The following analysis is similar to Scenario 1. However, after the addition of auxiliary qubit $|0\rangle_{aux}$, The unitary matrices implemented by Alice and Charlie differs from scenario 1. In Table 4, we have listed the unitary

Table 4 The unitary matrices of Alice and Charlie and success probability of three cases

Possible entangled states	A and C's unitary matrices	Success probability
$|\phi^+_{\theta'}\rangle_{A_1C_1} \otimes |\phi^+_{\theta'}\rangle_{A_2C_2}$	$U_2^{A_2aux}, U_2^{C_1aux}$	$\frac{4cos^4\theta}{(sin^2\theta+cos^2\theta)^2}$
$|\phi^+_{\theta^*}\rangle_{A_1C_1} \otimes |\phi^+_{\theta'}\rangle_{A_2C_2}$	$U_2^{A_2aux}, U_3^{C_1aux}$	$\frac{4cos^8\theta}{(sin^2\theta+cos^2\theta)(sin^6\theta+cos^6\theta)}$
$|\phi^+_{\theta'}\rangle_{A_1C_1} \otimes |\phi^+_{\theta^*}\rangle_{A_2C_2}$	$U_3^{A_2aux}, U_2^{C_1aux}$	$\frac{4cos^8\theta}{(sin^2\theta+cos^2\theta)(sin^6\theta+cos^6\theta)}$
$|\phi^+_{\theta^*}\rangle_{A_1C_1} \otimes |\phi^+_{\theta^*}\rangle_{A_2C_2}$	$U_3^{A_2aux}, U_3^{C_1aux}$	$\frac{4cos^{12}\theta}{(sin^6\theta+cos^6\theta)^2}$

matrices implemented on $A_2(C_1)|0\rangle_{aux}$ and the success probability of three cases. Where,

$$U_2 = \frac{1}{sin\theta}\begin{pmatrix} cos\theta & \sqrt{sin^2\theta - cos^2\theta} & 0 & 0 \\ 0 & 0 & 0 & sin\theta \\ 0 & 0 & sin\theta & 0 \\ \sqrt{sin^2\theta - cos^2\theta} & -cos\theta & 0 & 0 \end{pmatrix} \tag{16}$$

$$U_3 = \frac{1}{sin^3\theta}\begin{pmatrix} cos^3\theta & \sqrt{sin^6\theta - cos^6\theta} & 0 & 0 \\ 0 & 0 & 0 & sin^3\theta \\ 0 & 0 & sin^3\theta & 0 \\ \sqrt{sin^6\theta - cos^6\theta} & -cos^3\theta & 0 & 0 \end{pmatrix} \tag{17}$$

4 Conclusion

This work presented a quantum dialogue protocol for satellite-to-ground networks, which utilizes partially entangled states as quantum resources. In our protocol, any two communication clients can establish quantum dialogue by SGS's entanglement swapping in this networks. After the Bell-basis measurements, they can recover the target qubits each other by adding auxiliary qubit $|0\rangle_{aux}$, implementing specially defined unitary matrices and performing appropriate Pauli matrices. At last, we calculated the success probability of each scenario.

Acknowledgments Thanks to the support by the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20135134110003), Scientific Research Fund of Sichuan Provincial Education Department (No. 11ZB153) and Scientific Research Fund of Sichuan Minzu College (No. 13XYZB011 and No. 12XYZB006).

References

1. Nielsen, M.A., Chuang, I.L.: Cambridge University Press, Cambridge (2000)
2. Nguyen, B.A.: Phys. Lett. A **328**, 6–10 (2004)
3. Shen, D.S., Ma, W.P., Yin, X.R., Li, X.P.: Int. J. Theor. Phys. **52**, 1825–1835 (2013)
4. Wang, M.L., Ma, W.P., Shen, D.S., Wang, L.L.: Int. J. Theor. Phys. **54**, 1388–1395 (2015)
5. Sheng, Y.B., Zhou, L., Zhao, S.M., Zheng, B.Y.: Phys. Rev. A **85**, 012307 (2012)
6. Peng, J.Y., Bai, M.Q., Mo, Z.W.: Chin. Phys. B **23**, 10304 (2014)
7. Dong, L., Xiu, X.M., Gao, Y.J., Chi, F.: Opt. Commun. **281**, 6135–6138 (2008)
8. Yu, X.T., Xu, J., Zhang, Z.C.: Acta Phys. Sin. **61**, 220303 (2012)
9. Yu, X.T., Xu, J., Zhang, Z.C.: Chin. Phys. B **22**, 090311 (2013)
10. Wang, K., Yu, X.T., Lu, S.L., Gong, Y.X.: Phys. Rev. A **89**, 022329 (2014)
11. Chen, N., Quan, D.X., Pei, C.X., Yang, H.: Chin. Phys. B Sci. **24**, 020304 (2015)

A Rain Detection and Removal Algorithm Based on Rainy Intensity for Videos in Heavy Rainy Scene

Chao-qi Ma and Zheng-fa Liang

Abstract Detection and removal of rain for videos is a challenging problem because of the difficulty to build an appropriate imaging model for rainy weather. By far, almost all the algorithms proposed for rain removal considered a certain pixel in a video frame as rain affected pixel or non-rain affected pixel determinedly, which is not reasonable for heavy rainy scene. Because in the heavy rainy scene, almost all the pixels are covered by rain more or less. Instead of strictly considering a pixel as either rain covered or non-rain covered, the proposed algorithm calculates the rainy intensity for each pixel. The rainy intensity is soft information ranging from zero to one which contains more information about the rain. Furthermore, the rainy intensity can help estimate the original value of pixels that are covered by rain. Experiment results show that the rain removal videos processed by the proposed algorithm has much better vision effect than by other existing algorithms in heavy rainy scene.

Keywords Rain detection · Rain removal · Rainy intensity · Heavy rainy scene

1 Introduction

Weather conditions have great influence on the imaging quality of vision systems, thus making many applications of video analysis and computer vision such as motion detection, object tracking, video surveillance, and robot navigation and so on, lapse. It is a meaningful work to improve the image quality of videos degraded by bad weather conditions.

In general, weather conditions are broadly classified as steady (fog, mist and haze) or dynamic (rain and snow) according to the sizes of individual droplets [1, 2]. For the steady weather condition, taking haze removal for example, researchers have developed lots of algorithms based on the atmosphere scatter model, and have gotten very good results [3]. But there is no any imaging model for the rain removal

C. Ma (✉) · Z. Liang
National University of Defense Technology, Changsha 410073, China
e-mail: machaoqi1990@hotmail.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_61

problem. The common assumption is that the image degraded by rainy weather is composed of the original image and the rain drop. So the task is to detect all the regions covered by rain for every frame in a video, and then use the spatial or temporal information to recover the original value of the rain affected pixels. Based on this assumption, many good algorithms have been proposed.

Garg and Nayar [4] are the first researchers dealing with the rain removal problem. They analyzed the physical and photometric properties of the rain, finding that there is a positive intensity change when the pixel is covered by rain, and the rain drops in a video have the uniform velocities and directions.

Zhang et al. [5] proposed an algorithm that incorporates both temporal and chromatic properties of rain in video. The temporal property states that an image pixel is never always covered by rain throughout the entire video. The chromatic property states that the changes of R, G, and B values of rain affected pixels are approximately the same. For rain removal, they used the entire video to compute its intensity histogram of the pixel, and then used K-means clustering with $K = 2$ to divide the pixel to background and rain. The new color of a pixel affected by rain is replaced by the -blending of its rain-affected color and background color.

The models proposed in early years cannot work well in many occasions, such as highly dynamic scene, heavy rainy scene, and give limited performance. In recent years, some robust algorithms have been proposed.

Tripathi [6] proposed a probabilistic spatiotemporal approach for detection and removal of rain. They found that the time evolution of intensity values of a pixel at particular position present in rain region for consecutive frames is quite different from the evolution of pixel present in moving object region. And they used two statistical features, i.e. intensity fluctuation range, and spread asymmetry collected from spatial-temporal neighborhood to detect the rain affected pixels. But this method still gives lots of false detections.

Zhao [7] proposed a pixel-wise framework combining a detection method with a removal approach. In their work, dynamic weather conditions are detected by a strategy-driven state transition, which integrates static initialization using K-means clustering with dynamic maintenance of Gaussian mixture model. Moreover, a variable time window is presented for removal of rain and snow. However, this method is sensitive to fast intensity change and is of high computing complexity.

Chen [8] proposed rain pixels recovery algorithm based on motion segmentation scheme, that each pixels dynamic property as well as motion occlusion clue is considered; both spatial and temporal information are adaptively exploited during rain pixel recovery. This method can deal with rainy scenes with large motion very well. But the rain detection in this method is too simple, and it still cannot give perfect result for heavy rainy scene.

Zhou et al. [9] proposed a rain removal algorithm based on optical flow and hybrid properties constraint. This approach firstly identified the candidate rain pixels by adopt the optical flow, and then used hybrid properties constraint of raindrop to refine the rain streaks. Once the rain streaks were detected, the scene can be restored by a weighted composing method.

The common problem shared by the above algorithms is that they considered a pixel as rain affected or not determinedly. It is not reasonable especially for heavy rainy scene. In a heavy rainy scene, almost all the regions are covered by rain more or less, that is the rain is lighter in some regions, and heavier in other regions. We should establish a model to estimate the rain heavy level of each pixel in a frame. The advantage of this model is that all the pixels including rain pixels and non-rain pixels, can be deal with in a unique framework. As a special case, we can simply consider the rain heavy level of background pixels as zero. And in the rain removal stage, we can estimate the original value of the rain affected pixels with the help of this model naturally.

In this paper, we proposed a model to estimate the rain heavy level with two features extracted from the time evolution information of a pixel. The rain heavy level is called rainy intensity. We calculate each pixels rainy intensity for each frame in a video, instead of detecting its pixels as rain affected or not. The rainy intensity is later used to help estimate the original value of a pixel. Experiment results show that the proposed algorithm performs better than existed algorithms in heavy rainy scene.

The rest of the paper is organized as follows. Section 2 describes the models establishing process based on analyzing the rains characters. Section 3 will show the rain removal framework based on the proposed model. In Sect. 4, we compare our algorithm with some existed excellent algorithms.

2 Rain Analysis and Rainy Intensity

In our work, we suppose that the camera is static and the video is stable. If the video is twittering, it always can be stabilized using algorithms like video stabilization [10]. And we suppose the rain drops are randomly distributed in the space just as many other researchers do.

In a rainy video, a main phenomenon of a pixel affected by rain is temporal fluctuations [4], that is, rain gives positive fluctuations in the intensity values without affecting the chrominance values. Here, we first study the time evolution for pixels of rain affected, background and motion objects. Then for the rain affected pixels, we consider the difference between heavy rain and light rain.

Figure 1 shows the time evolution of pixels that belong to rain, background, and motion object in the 6th frame (Fig. 1a) of 11 continuous frames respectively in a highly dynamic scene. For each situation, we take three pixels into consideration. The precise values of these pixels are listed in Table 1.

It is obviously that if a pixel is affected by rain, there is a positive intensity fluctuation along the pixels time evolution and its former and latter intensity values are almost symmetric. If the pixel belongs to background, almost no changes in intensity can be detected. And for motion objects, there is no regular pattern can be found.

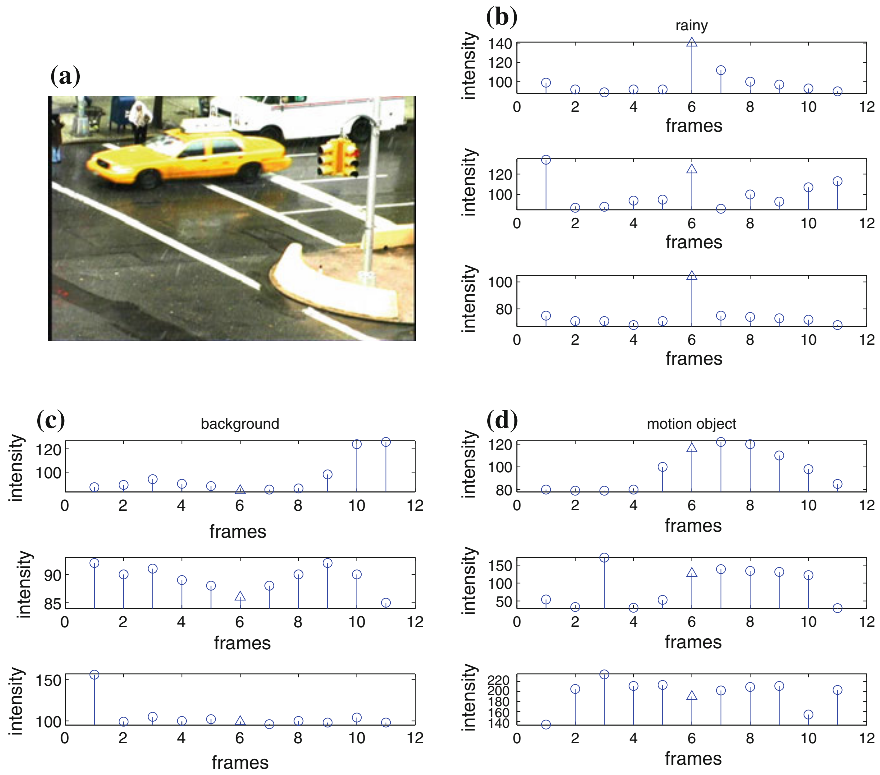


Fig. 1 The time evolution of pixels intensity. **a** The current frame. **b** Pixels affected by rain. **c** Pixels belong to background. **d** Pixels covered by motion object

Table 1 The specific values of the pixels in Fig. 1

	$N-2$	$N-1$	N	$N+1$	$N+2$
Rainy	92	92	140	112	100
	94	95	124	86	100
	68	71	104	75	74
Background	90	88	84	85	86
	89	88	86	88	90
	100	102	99	96	100
Motion object	80	100	116	122	120
	217	190	235	187	173
	211	213	190	202	209

Similar results can be found in A.K. Tripathi's work [6]. But they thought that the symmetric feature can be still retain when consider the pixels spatial neighbors thus using a spatiotemporal window to measure the pixels symmetric property. In fact, in the spatial domain, motion objects can present better symmetric property because

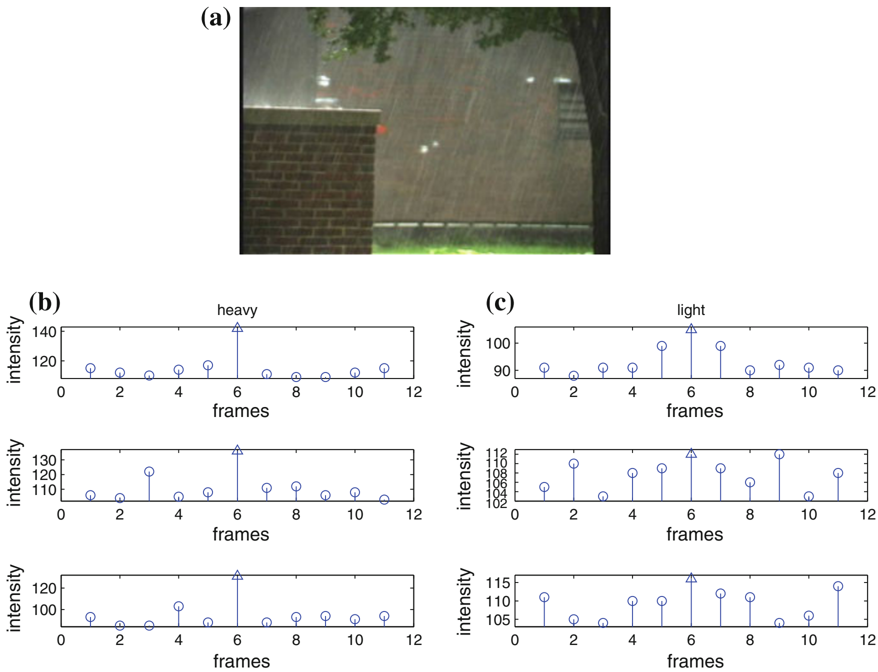


Fig. 2 The time evolution of pixels intensity. **a** The current frame. **b** Pixels affected by light rain. **c** Pixels affected by heavy rain

usually objects have the same color in large continuous regions. So in our work, we will not include spatial information to measure one pixels symmetric property, but only consider its temporal information.

Next, we study the difference between heavy rain and light rain in a heavy rainy scene. In actual world, when its rainy outside, rain drops exist in different visual depth. So a more reasonable assumption is that all the pixels in a frame are covered by rain, and for individual pixels, their rain intensity is different. Figure 2 shows the differences in pixels intensity which are covered by light rain and heavy rain respectively.

Table 2 gives the specific values of the pixels in Fig. 2. We found that the positive fluctuation of pixels affected by heavy rain is much bigger than that of pixels affected by light rain. So we propose two features extracted from continuous five frames to distinguish rain, background, and motion object, furthermore to distinguish heavy rain and light rain. The two features are named average positive intensity fluctuation $A(x)$, and symmetric level $B(x)$.

Table 2 The specific values of the pixels in Fig. 1

	$N-2$	$N-1$	N	$N+1$	$N+2$
Light	91	99	105	99	90
	108	109	112	109	106
	110	110	116	112	111
Heavy	114	117	142	111	109
	105	108	136	111	112
	103	88	131	88	93

$$A(x) = \begin{cases} \dfrac{mean\{d_n\}}{A_0}, & mean\{d_n\} \geq 0, n = -2, -1, 1, 2 \\ 0, & other \end{cases} \quad (1)$$

$$B(x) = \begin{cases} \dfrac{2d_{-1}d_1}{d_{-1}^2 + d_1^2 + d_0} \cdot \dfrac{2d_{-2}d_2}{d_{-2}^2 + d_2^2 + d_0}, & d_n > 0, n = -2, -1, 1, 2 \\ 0, & others \end{cases} \quad (2)$$

$$d_{-2} = I_N - I_{N-2}, d_{-1} = I_N - I_{N-1}, d_1 = I_N - I_{N+1}, d_2 = I_N - I_{N+2}$$

Here, A_0 is a constant chosen to guarantee $A(x) \in [0, 1]$. From the definition of $B(x)$ we can easily find that $B(x) \in [0, 1]$. $A(x)$ and $B(x)$ measure the positive fluctuation level and symmetric property of the current pixel over its temporal neighbors, respectively. The higher value means high positive fluctuation or high symmetric, which means the current pixel is more likely to be covered by rain and the rain heavy level is higher. We call the rain heavy level rainy intensity. One possible way to calculate the rainy intensity of a pixel can be:

$$p(x) = \frac{1 - e^{-f(A)g(B)}}{1 - e^{-1}} \quad (3)$$

Here, $f(x)$ and $g(x)$ are increase functions of variable x, and satisfy:

$$f(x) \in [0, 1], g(x) \in [0, 1], for\ x \in [0, 1]$$

For simplicity, we take

$$\begin{cases} f(x) = x \\ g(x) = x \\ p(x) = \dfrac{1 - e^{-AB}}{1 - e^{-1}} \end{cases} \quad (4)$$

This model is simple, but it works.

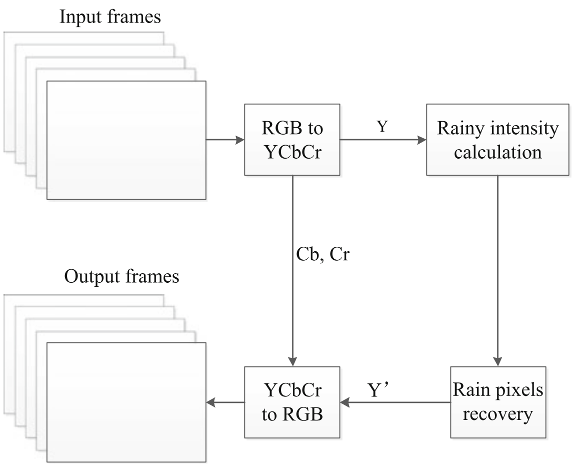


Fig. 3 Main framework of the proposed algorithm for rain removal

3 Rain Removal Based on Rainy Intensity

In this section, we will detailedly describe the rain removal method using the proposed rainy intensity.

The framework of the proposed algorithm is shown in Fig. 3. For the frames in the video taken in rainy weather, RGB to YCbCr conversion is first done so that the rain removal work can be exploited only on the intensity plane. After that, rainy intensity calculation is done to prepare the necessary information for rain pixels recovery. The recovered pixels then form a new intensity plane for current frame. At last, the new intensity plane together with the Cb and Cr component make up the final rain removal frame.

3.1 Rain Intensity Calculation

The rainy intensity of a frame is calculation based on the Eqs. (2)–(4) in Sect. 2. In our method, only five continuous frames and their corresponding rainy intensity are stored in a FIFO to reduce system cost. The unique frame work to calculate the rainy intensity, as well as the recovery method described next, make it easier to apply this algorithm to VLSI implementation.

A problem necessary to discuss is that the method proposed sometime give high rainy intensity value to large motion objects. In addition, motion objects may also be covered by rain drops, and the model is more complicate. In fact, all the algorithms working on pixel level cannot deal with this problem very well. We need some method working on semantic level to distinguish motion objects from the current frame. In [8], motion segmentation is applied and it seems that the results are very well. But the motion segmentation method proposed in [8] is very complicated which involves optical flow, Gaussian Mixture Model (GMM), and image segmentation using k-means clustering. So the algorithm in [8] cannot achieve real time processing. Developing a new simple method to deal with this problem then becomes very meaningful and it is our future work as well.

3.2 Rain Pixels Recovery

After calculating the rainy intensity for every pixel in the current frame, the rain pixels recovery job can be done.

According to the researches done before, the original value of the rain-affected pixel could be estimated by a weighted sum of the pixels from both temporal and spatial neighborhood. But experiments show that using spatial neighborhood to recover the original value of a pixel can blur the frame. So in our work we only use their temporal information to estimate the original value.

Different from other pioneer works, we assume that all the pixels in rainy scene are covered by rain more or less. We cant just simply use the temporal mean, or weighted temporal mean with fixed coefficient. An appropriate way is to adjust the coefficients according to their heavy level, which can be measured by the rainy intensity proposed ahead. So one possible way to estimate the rain affected pixels original value can be:

$$\hat{I}(x) = \frac{\sum_{n=-2}^{2} \left[1 - p_n(x)\right] \cdot I_n(x)}{\sum_{n=-2}^{2} \left[1 - p_n(x)\right]} \tag{5}$$

4 Experiment Results

Experiments are carried out on two videos of heavy rain scene and light rainy scene in *Matlab 2012a* environment. The results are shown in Figs. 4, 5 and 6.

The rainy intensity calculated for these two videos are shown in Fig. 4. From the results we can see that the proposed rainy intensity can measure the heavy level of rain very well.

Figures 5 and 6 show the rain removal results of two state of art algorithms, and compare them with our algorithm. The two algorithms were proposed by Tripathi [6] in 2012, and Chen [8] in 2014 respectively. From the results we can see that the rain streaks are well removed in these two videos. And the visual effect of the frame processed by our algorithm is better than that by other algorithms.

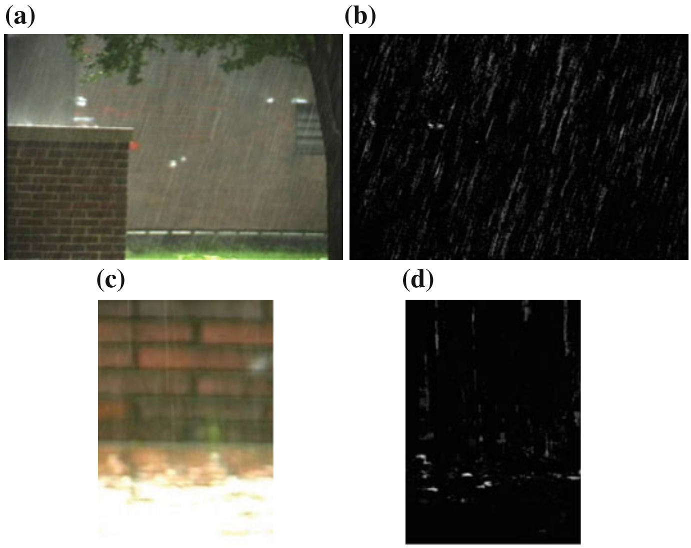


Fig. 4 Simulation results of two video showing the original frames and their corresponding rainy intensity

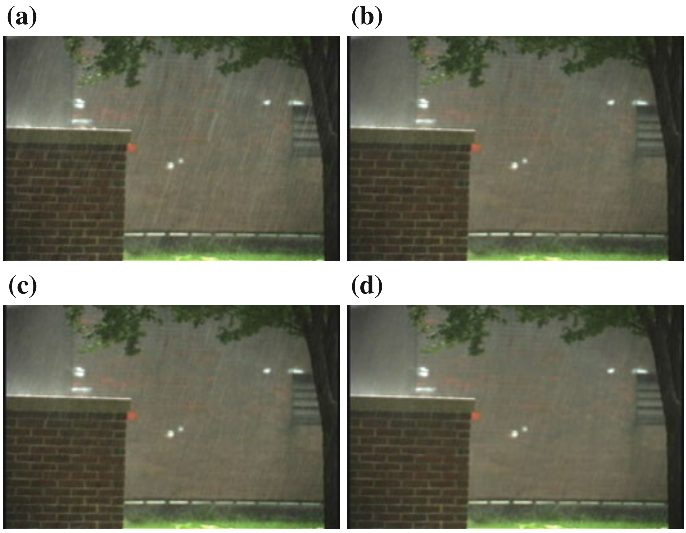


Fig. 5 Comparison on the first video using different rain removal algorithms. **a** Original video frame. **b** A. Tripathi's algorithm using spatial-temporal model. **c** Jie Chens algorithm using motion segmentation and adaptive filters. **d** Result from our proposed algorithm

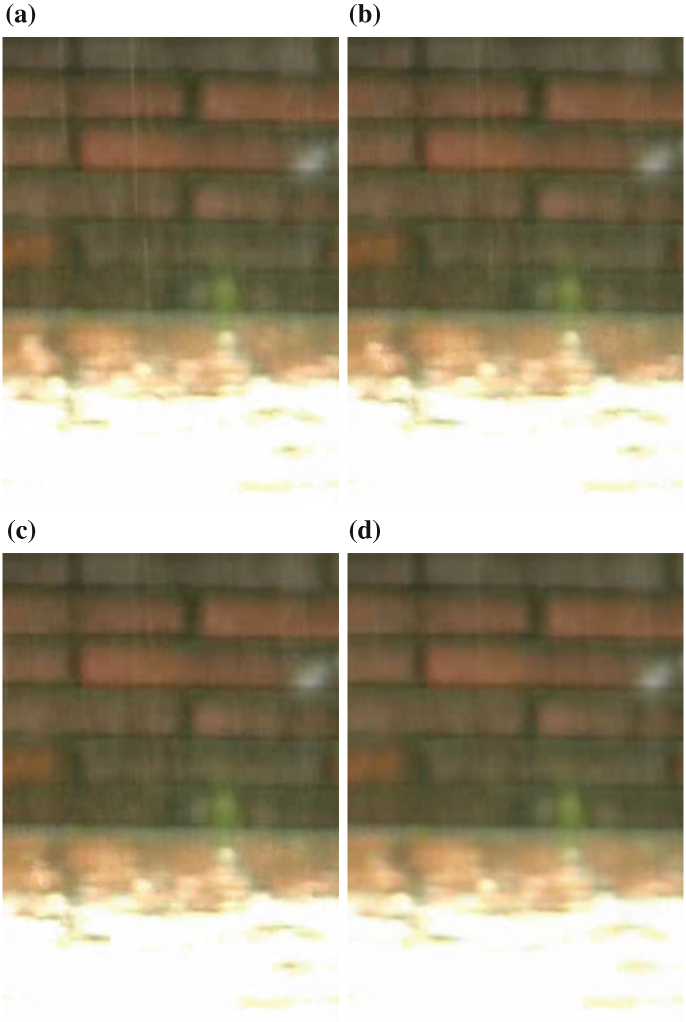


Fig. 6 Comparison on the second video using different rain removal algorithms. **a** Original video frame. **b** A. Tripathi's algorithm using spatial-temporal model. **c** Jie Chens algorithm using motion segmentation and adaptive filters. **d** Result from our proposed algorithm

5 Summary

In this paper, we present a novel method to remove rain streaks in videos based on rainy intensity proposed. The rainy intensity is defined according to the positive intensity fluctuations and intensity symmetric along the pixels time evolution. The rainy intensity measures the rains heavy level very well, and helps to estimate the pixels original value in highly rainy scene. Experiments show that our algorithm can deal with both light and heavy rainy scenes with lower computation. And most importantly the visual effect of video after rain removal by our algorithm is much better than existed algorithms.

References

1. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vis. **48**(3), 233–254 (2002)
2. Garg, K., Nayar, S.K.: Vision and rain. Int. J. Comput. Vis. **75**(1), 3–27 (2007)
3. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2011)
4. Garg, K., Nayar, S.K.: Detection and removal of rain from videos. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2004, vol. 1, pp. I–528. IEEE (2004)
5. Zhang, X., Li, H., Qi, Y., Leow, W.K., Ng, T.K.: Rain removal in video by combining temporal and chromatic properties. In: 2006 IEEE International Conference on Multimedia and Expo, pp. 461–464. IEEE (2006)
6. Tripathi, A.K., Mukhopadhyay, S.: Video post processing: low-latency spatiotemporal approach for detection and removal of rain. IET Image Process. **6**(2), 181–196 (2012)
7. Zhao, X., Liu, P., Liu, J., Tang, X.: Removal of dynamic weather conditions based on variable time window. Comput. Vis., IET **7**(4), 219–226 (2013)
8. Chen, J., Chau, L.-P.: A rain pixel recovery algorithm for videos with highly dynamic scenes. IEEE Trans. Image Proc. **23**(3), 1097–1104 (2014)
9. Zhou, Y., Han, Y., Zhou, P.: Rain removal in videos based on optical flow and hybrid properties constraint. In: 2015 Seventh International Conference on Advanced Computational Intelligence (ICACI), pp. 143–147. IEEE (2015)
10. Matsushita, Y., Ofek, E., Tang, X., Shum, H.-Y.: Full-frame video stabilization. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2005, vol. 1, pp. 50–57. IEEE (2005)

Attribute Classification for Transformer Substation Based on Deep Convolutional Network

Jia Wu, Dan Su, Hong-fei Xu, Si-ri Pang and Wang Luo

Abstract Search engines of transformer substation receive a large number of category-related and content-based queries. These queries are best answered by the attribute listings, which contain the semantic object information. While creating such listing needs the requirement analysis to figure out the most concerned attributes. In this paper, we build a new dataset of transformer substation for attribute classification. To efficiently annotate the image set, we employ the deep convolutional network, which outputs the posterior probability of each attribute. The extensive experiments show that our work reaches 94.67 % for attribute classification, which demonstrate the efficiency of proposed work.

Keywords Attribute listings · Attribute classification · Deep convolutional network

1 Introduction

With the increase of the monitoring devices of the transformer substation, the monitoring images and videos increase rapidly. Due to the time consuming of hand annotations, users need a search tool to automatically find the query-related images.

J. Wu (✉) · D. Su · H. Xu · S. Pang
State Grid JIBEI Electric Power Company, Beijing, China
e-mail: 13811810276@163.com

D. Su
e-mail: su.dan@jibei.sgcc.com.cn

H. Xu
e-mail: xu.hongfei@jibei.sgcc.com.cn

S. Pang
e-mail: pang.siri@jibei.sgcc.com.cn

W. Luo
NARI Group Corporation, Nanjing, China
e-mail: luowang@sgepri.sgcc.com.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_62

Fig. 1 Three images of transformer substation containing “person”. The *red box* indicates the location of person. A large variation of pose, appearance and scale is observed in these images

These queries can be answered based on the attribute listings, which contain the middle-level semantic information.

Creating an exact and useful attribute list is a difficult and time consuming work, which requires the investigation of the requirements of users. The list should contain the category-related and object-level information. With the augment of the monitoring images and videos, their categories increase rapidly which add the difficulty of designing an attribute list.

After the creation of the attribute list, a dataset of transformer substation is required to train the automatic method of attribute annotation. This dataset should contain a variety of images, taken from different angle, illumination and scale.

In this paper, we focus on the attribute classification for transformer substation. Given a query image, our task is to generate the corresponding attribute listing. This is a challenging task since there exists a large variation of the appearance of the object. For instance, we show three images containing attribute “person” in Fig. 1. The variations of clothes, illumination and scale make it difficult to recognize this attribute.

In our work, we introduce a new dataset of transformer substation. These images are collected from the photos and internet. The variation of background, scale and angle make this dataset difficult for attribute classification. Based on this dataset, we train a deep convolution network to generate the attribute listings for query images. To avoid the overfitting, we transfer the parameters of the pre-trained model on the large-scale dataset to our deep convolutional network. The last layer is replaced with sigmoid layer for multi-label classification. The extensive experiments show that our method reaches 94.67 % for 7 attribute classification, which demonstrates the efficiency of our method.

This paper is organized in the following way. In Sect. 2, we review previous related works on classification and neural network. In Sect. 3, the details of proposed method and dataset are introduced. Section 4 describes the experimental setup and presents the results of experiments. Finally, we conclude our work in Sect. 5.

2 Related Word

Attribute classification is a useful area in computer vision, which can be treated as the pre-processing of many tasks (image classification [1] and fine-grained recognition [2, 3]). Here, we focus on the reviewing of related works on image classification and the structure of deep convolutional network.

Image classification. Traditional works [4] on the image classification follows the pipeline of low-level feature extraction, feature coding and classifier training. In the first step, the SIFT [5] or HOG features [6] are extracted from the dense patches of the images. Then, the middle-level features are obtained based on Bag-of-Words (BoW) representation [4]. Finally, these features are fed to the classifiers such as SVM [7] or random forest [8]. Since BoW representation ignores the spatial relationship of local features, [9] introduces the spatial pyramid which partitions the image into several regions. The concatenated histogram of all regions is employed to represent the image.

Deep Convolutional Network. In recent years, convolutional network [10] has led to the state-of-arts performance for many tasks: image recognition and object detection. The convolutional network is a kind of neural network with millions of parameters, which is robust to some extent of variations of scale and illumination. Different from the hand-crafted feature, the convolutional network learns the classifier and features iteratively.

Due to the large amount of parameters, avoiding overfitting is a serious problem. To achieve this, convolutional network use the large-scale dataset [11] which has more than one million images to train the model. To train a small dataset specific network, one can transfer the parameters pre-trained by the large dataset and apply finetune processing [12]. We base our work on this infrastructure.

3 The Proposed Method

In this section, we first introduce the dataset of transformer substation. The difficulties of this dataset are highlighted. Then, the details of the deep convolutional network are described.

3.1 Dataset of Transformer Substation

As the first step to solve the problem of attribute classification, there is no existing dataset of transformer substation. Instead, we introduce a new dataset containing 7 attributes collected from the internet and photos: transformer substation, transformer, disconnecting link, transmission line, insulator, business hall and person.

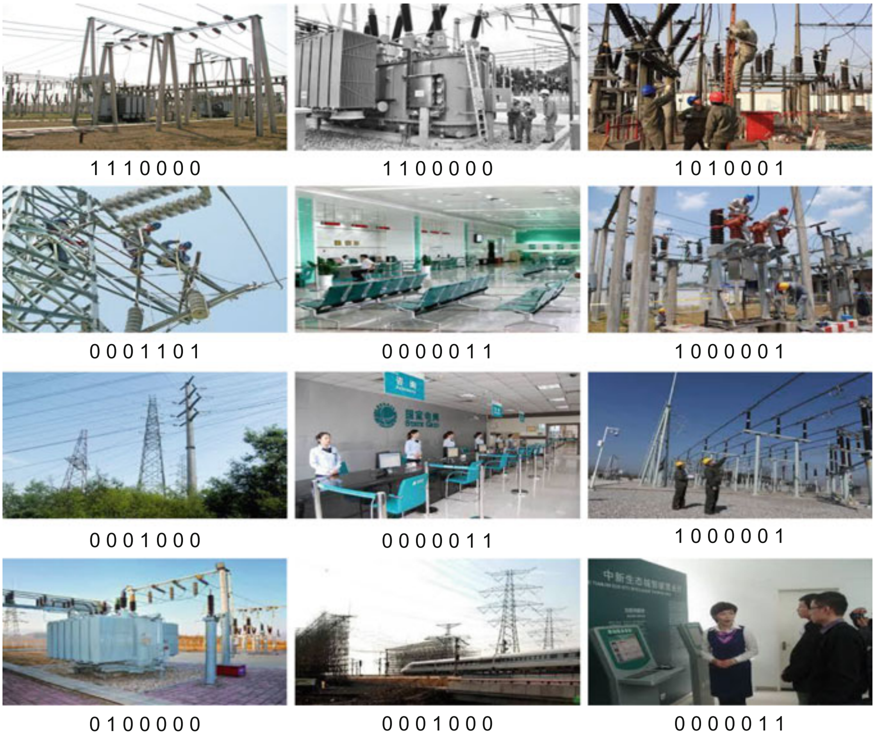


Fig. 2 Our dataset contains 7 attributes: transformer substation, transformer, disconnecting link, transmission line, insulator, business hall and person. We give a attribute listing shown below each image

The numbers of images for each attribute ranges from 60 to 140. Each image is labeled with the attribute listing. For instance, the listing {1, 1, 1, 0, 0, 0, 0} below the first image in Fig. 2 indicates that there are transformer substation, transformer and dis connecting link in this image. Based on these images, we can see that this attribute dataset of transformer substation is very challenging and difficult. Here, we highlight several difficulties of this dataset:

- There is a large variation of pose, illumination, scale and angle.
- The background of the image is cluttered.
- Some objects are very small. It is difficult to find them.
- The image number varies a lot for different attributes.

3.2 Deep Convolutional Network for Attribute Classification

Our convolutional network is based on the framework of Alex convolutional network [10]. This network has drawn many attentions due to its impressive perform in large-scale image classification challenge. This work uses Convolution layer, Normalization layer, Pooling layer and Inner Product layer to capture the semantic information of the image. The network use backward propagation to update the parameters within each node of the layer. Since traditional CNN model is used for single-label image classification, which outputs the label with the highest probability. While, our work focus on attribute classification, which annotates an image with multi-label. To solve this problem, we replace the last layer (Soft max layer) with sigmoid layer. The flowchart of neural network is shown in Fig. 3. When the probability of the attribute is higher than 0.5, we assign 1 for this attribute. To train a network optimized for multi-label classification, we define our loss function as follows:

$$L_{loss} = \sum_{k=1}^{K} (l_k \log p_k + (1 - l_k) \log(1 - p_k)) \tag{1}$$

where L_{loss} is the loss energy, l_k is the ground-truth label of k-th attribute, and p_k indicates the corresponding posterior probability. We use forward and backward propagation to iteratively update the model parameters.

We split our dataset into train/test data randomly. Since a convolutional network contains millions of parameters, using small dataset to train the network will cause a serious overfitting. To solve this problem, we pre-train the convolutional network

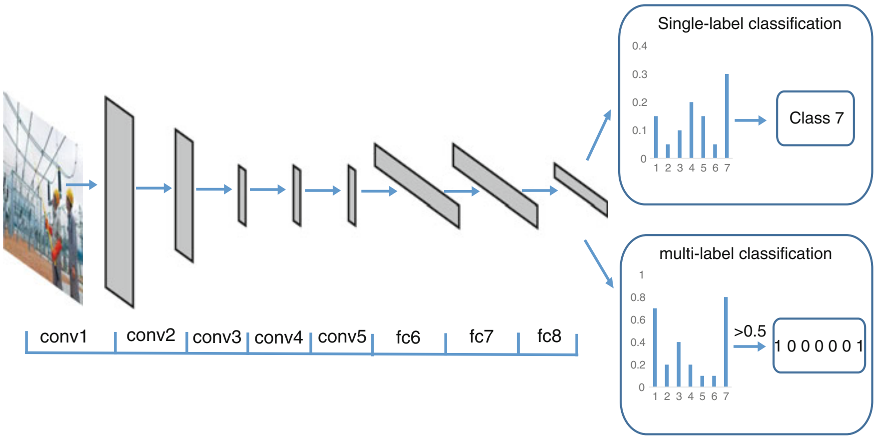


Fig. 3 The deep convolutional network for single-label and multi-label image classification. In the bottom, we show the name of 7 layers in this network: conv1, conv2, conv3, conv4, conv5, fc6, fc7, where"conv" is short of"convolution" and"fc" is short of"full connected"

using the large-scale ImageNet dataset. Then, the final Inner Product and Soft max layer are replaced. Since we focus on attribute classification, each image is annotated with multi-label. The final layer is replaced with Sigmoid layer. Finally, we continue to train the model using our dataset of transformer substation until convergence.

Each image is resized to 256 × 256 pixels. During the training period, we randomly generate 224 × 224 crops from each training image. The input images are normalized based on the mean value of images. The mirrored versions of training images are added to increase training data. To avoid overfitting, we set the dropout as 0.7, which means 70 % parameters are randomly chosen to be not used.

4 Experiment

In this section, we describe the experimental setup and show the results of attribute classification. To start with, we analyze the performance of the proposed method. Then, we show what features can be leaned by deep convolutional network and concluded with quantitative results.

4.1 Experimental Setup

Our experiments are evaluated at the introduced dataset of transformer substation. Each image is annotated with binary listing indicates whether there is the corresponding object in the image. We split the dataset into training and testing sets: 50 % images for training and the rest images for testing.

For pre-trained convolutional network, the dimension of fc8 is 1000, since there are total 1000 classes in this largescale dataset. For our dataset, there are total 7 attributes, so we replace fc8 with the Inner Product layer with 7 outputs, followed by the sigmoid layer. To optimize this network, we optimize the parameters based on the sigmoid loss function defined by Eq. 1. The backward propagation [10] is employed to calculate the gradient of parameters. Then, gradient descent optimization [13] method is used to update the parameters. Our work is based on caffe framework [14]. The learning rate is initially set as 0:000001.

4.2 Results of Evaluation

During the test period, the query image is resized to 224 × 224 pixels. Then, it is fed to the deep convolutional network to output the probability distribution for all

Table 1 Attribute classification

Attribute	Accuracy (%)
Transformer substation	92.54
Transformer	93.42
Disconnecting link	91.67
Transmission line	96.93
Insulator	93.86
Business hall	100
Person	94.30
Mean	94.67

Fig. 4 Four examples for person attribute classification. We used *black box* to overlap the region of person. The probability of person is shown below the image

attributes. Based on the threshold 0.5, the predicted attribute listing is obtained. Compared with the ground-truth listings, we compute the accuracy for each attribute. The results of experiments are shown in Table 1. It can be seen that our method obtains the accuracy 94.67 %. We show some results of attribute classification in Fig. 4. It observes that our method can obtain the accurate results for images with a large variation of scale, illumination and appearance.

We investigate what can be learnt by the deep convolutional network. As shown in Fig. 5, the probability of person drops dramatically when the region of person is overlapped with a black box. This illustrates that the network can extract the object-level feature in a semi-supervised way.

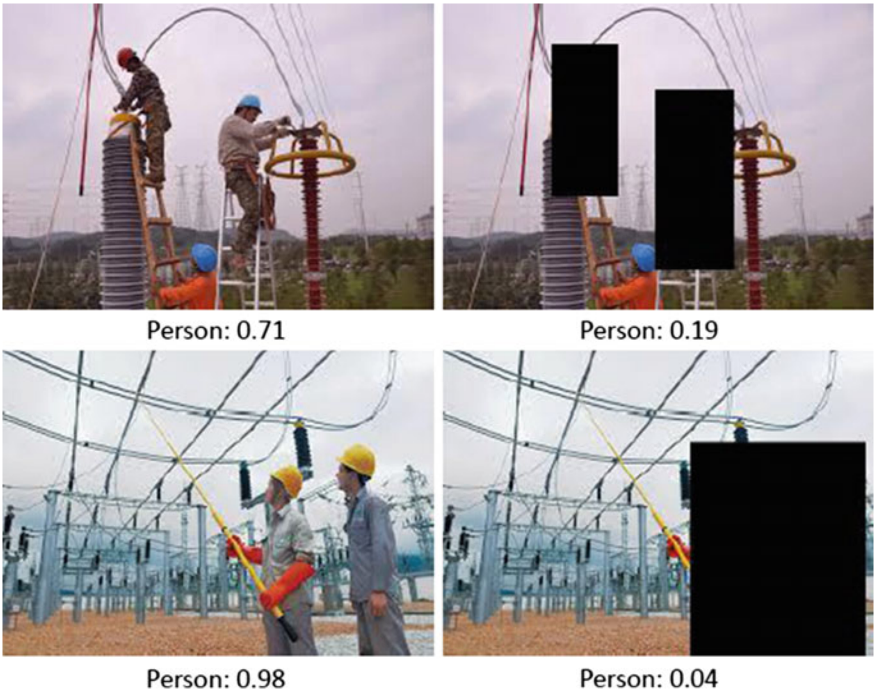


Fig. 5 Two results of attribute classification. We list four top ranked attributes below the image with the corresponding posterior probability

5 Conclusion

In this paper, we introduce a new dataset of transformer substation for attribute classification. A deep convolutional network is employed to generate the attribute listings. The last layer of traditional network is replaced by sigmoid layer to product multi-label. The backward propagation and gradient descent optimization are employed to optimize the parameters of the network. The experiments demonstrate the impressive performance and show that the network can learn the object-level feature. In the future, we are interested in using deeper and wider network to figure out the relationship between different attributes.

Acknowledgements This work was supported in part by the Natural Science Foundation of Jiangsu Province (Grants No. BK20130107).

References

1. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. J. Comput. Vis. (2001)
2. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: CVPR (2014)
3. Yair, M.-A., Qian, Y., Martin, C.S., Vinay, S., Sacha, A., Liron, Y.: Ontological supervision for fine grained classification of street view storefronts. In: CVPR (2015)
4. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: ECCV (2004)
5. Lowe, D.: Object recognition from local scale-invariant features. In: ICCV (1999)
6. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR (2005)
7. Hearst, M.A., Dumais, S., Osman, E., Platt, J., Scholkopf, B.: Support vector machines. Intell. Syst. Appl. (1998)
8. Breiman, L.: Random forests. Machine Learning (2001)
9. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: CVPR (2006)
10. Krizhevsky, I.S.A., Hinton, G.: Imagenet classification with deep convolutional neural networks. In: NIPS (2012)
11. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: Decaf: a deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531 (2013)
12. Wei, Y., Xia, W., Huang, J., Ni, B., Dong, J., Zhao, Y., Yan, S.: Cnn: single-label to multi-label. arXiv preprint arXiv:1406.5726 (2014)
13. Bishop, C.M., Nasrabadi, N.M.: Pattern recognition and machine learning (2006)
14. Jia, Y.: Caffe: an open source convolutional architecture for fast feature embedding. In: http://caffe.berkeleyvision.org/ (2013)

Attribute Segmentation for Transformer Substation Using Convolutional Network

Jia Wu, Hong-zhen Yang, Hong-fei Xu, Si-ri Pang, Huan-yuan Li and Wang Luo

Abstract Convolutional network has shown its impressive performance in many tasks of computer vision: image classification, object recognition and detection. In this paper, we focus on the pixelwise attribute classification based on convolutional network. Different from traditional semantic segmentation which assigns single-label to a pixel, attribute segmentation is a multi-label segmentation problem. To solve this problem, we train the attribute classification network, followed by attribute-specific segmentation model. To reduce the spatial loss caused by the pooling layer, we propose a simple and efficient network with deconvolutional layer. Our method achieves a good performance in a new dataset of transformer substation.

Keywords Convolutional network · Attribute · Classification · Segmentation

J. Wu (✉) · H. Xu · S. Pang · H. Li
State Grid JIBEI Electric Power Company, Beijing, China
e-mail: 13811810276@163.com

H. Xu
e-mail: xu.hongfei@jibei.sgcc.com.cn

S. Pang
e-mail: pang.siri@jibei.sgcc.com.cn

H. Li
e-mail: li.huanyuan@jibei.sgcc.com.cn

H. Yang
State Grid ZHEJIANG Electric Power Company, Hangzhou, China
e-mail: yanghz@sina.com

W. Luo
NARI Group Corporation, Nanjing, China
e-mail: luowang@sgepri.sgcc.com.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_63

1 Introduction

In recent years, the convolutional network has achieved an outstanding performance compared with other state-of-arts methods in the large-scale image classification challenge. Many works based on convolutional network have been proposed to solve the problems of computer vision: object detection and semantic segmentation.

Compared with image classification which outputs a single label for the whole image, semantic segmentation assigns a label for each pixel of the image. This is a very difficult task, which suffers from a large variation of pose, scale and illumination. Previous methods on semantic segmentation mainly focus on assigning a single-label for each pixel. In fact, multi-label segmentation is more useful to reveal the attribute of the object. We show an example of attribute segmentation in Fig. 1. Some pixels in the region of person are also annotated by cloth or hat. Compared with single-label segmentation, multi-label segmentation is more attractive due to the rich information about the object.

In this paper, we focus on the attribute segmentation for transformer substation. Since there is no existing dataset for attribute annotation, we introduce the dataset of transformer substation, which gives the pixelwise annotation for each attribute. The convolutional network is employed to train the model for classification and

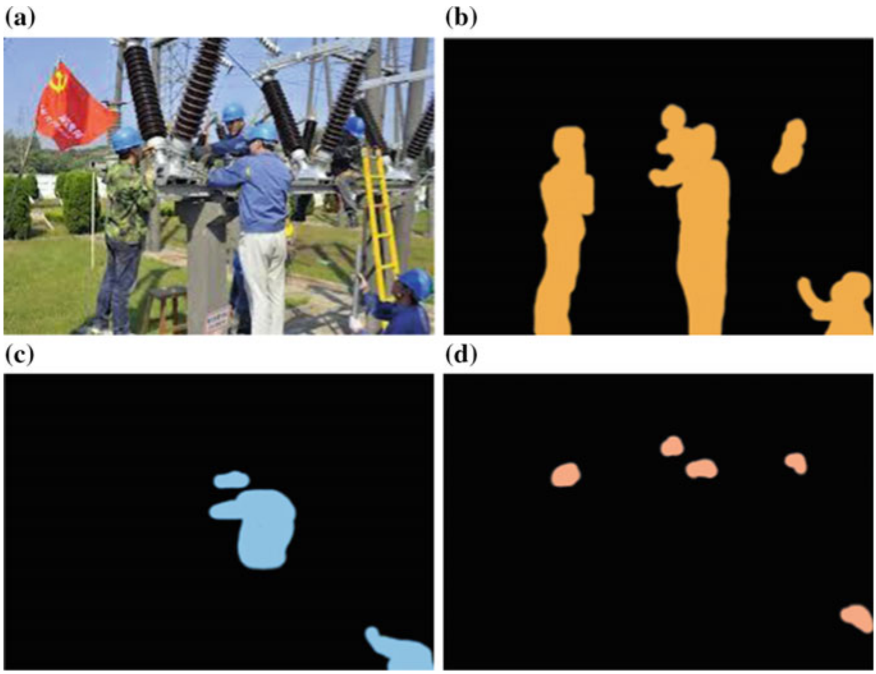


Fig. 1 The instance to show attribute segmentation which outputs multi-label for each pixel. **a** The input image. **b** The pixelwise annotation of attribute "person". **c** The segmentation of attribute "blue cloth". **d** The annotation of attribute "hat"

segmentation. To reduce the spatial loss of pooling layer, we propose a single and efficient convolutional network with deconvolutional layer. To obtain the attribute segmentation of the image, our work can be divided into two steps. The first step is to obtain the attribute list of the image. Then, the corresponding segmentation network of attribute is employed to automatically annotate the image.

In the next section, we review the related works on image classification based on deep network and recent approaches on semantic segmentation. The following section explains the structure of our network consisting of attribute classification with segmentation. Finally, we describe the details of experimental setup and the results of attribute segmentation, followed by the conclusion of our work.

2 Related Work

In recent years, deep convolutional work has been widely used in many areas of computer vision: image classification [1] and semantic segmentation [2]. Here, we review previous works on deep image classification and semantic segmentation based on the convolutional network.

Deep image classification The deep convolutional network has achieved the state-of-arts performance compared with other traditional methods based on the hand-crafted feature in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC12) [3] of image classification. The feature of the deep convolutional network is auto-learnt to optimize the defined loss function. There are two states to perform parameter optimization: forward propagation to compute the state of the neural units and backward propagation to calculate the gradients for each unit. The gradient descent method is employed to update the parameters. Due to the large amount of parameter number, it's a difficult problem to avoid the overfitting of the parameter. To tackle this problem, the large-scale dataset and data augmentation methods are employed to train the model.

Recent works train the convolutional network on some small dataset [2]. The parameters of pre-trained network using the large-scale dataset are transferred. Then, the network is fine-tuned to perform parameter optimization. To make the network capture more semantic information of the object, the deeper and wider network is a straightforward way. This will suffer from a large cost of memory and time. Motivated by this idea, a deep network named GoogleNet [4] is proposed with the inception module, which has achieved the best performance in ILSVRC14. The success of the network is based on its carefully designed structure which improves the utilization of the computing resources.

Semantic Segmentation Previous works on semantic segmentation mainly focus on outputting a single label for each pixel of the image. There are several traditional ways to produce the semantic annotation. Mostajabi et al. [5] generates the multiscale superpixels and classify each region to obtain the label. Dai et al. [6] performs the classification of region proposals and uses the map of segmentation to perform the refinement.

Table 1 The configuration of our network

Name	Kernel size	Stride	Pad	Output size
Input	–	–	–	224 × 224 × 3
conv1	11	4	100	104 × 104 × 96
conv2	5	1	2	52 × 52 × 256
conv3	3	1	1	26 × 26 × 384
conv4	3	1	1	26 × 26 × 384
conv5	3	1	1	26 × 26 × 256
conv6	6	1	0	8 × 8 × 4096
conv7	1	1	0	8 × 8 × 4096
conv8	1	1	0	8 × 8 × 2
deconv	63	32	0	287 × 287 × 2
crop	–	–	–	224 × 224 × 2

Here, "conv" indicates the convolutional layer and "deconv" indicates deconvolutional layer. The ReLu and LRN layer are omitted for brevity

In last three years, the convolutional network has shown its advantage in image classification and object detection which output a label for a whole image. A natural way is to produce a label for each pixel based on neural network. Recent works on convolutional network have driven a breakthrough for semantic segmentation. For example, fully connected network (FCN) [7] interpret the fully connected layers as convolutions with large receptive fields and perform image segmentation based on the coarse class score map. Noh et al. [8] uses the deconvolution and unpooling layers to reconstruct the label segmentation. This network adopts the convolutional layers adopted from the pre-trained networks on the large-scale dataset. Given a query image, the trained model is applied to each proposal of the image. The final result is obtained by combining the segmentation of all proposals (Table 1).

3 Proposed Method

This section introduces the dataset for attribute segmentation and describes the structure of the deep convolutional network.

3.1 Dataset for Attribute Segmentation

Different from recent works on semantic segmentation, which output a single label for each pixel, we focus on the attribute segmentation to produce pixelwise multi-label segmentation. To our knowledge, there is no existing dataset for attribute segmentation, we build a new dataset consisting 7 attributes: transformer substation, transformer, disconnecting link, transmission line, insulator, business

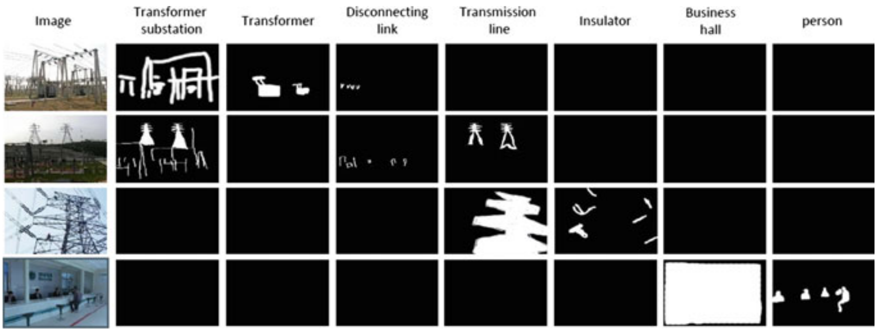


Fig. 2 The examples of dataset. We show the attribute map of each image

hall and person. For each image, we generate an attribute listing to indicate the existence of the corresponding attribute. The segmentation mask of each attribute is also annotated by hand. We show several examples in Fig. 2. It observes that one pixel may be assigned with multi-label.

3.2 *Network for Attribute Segmentation*

Our convolutional network is composed of two parts: the network for attribute classification and the convolutional network for attribute segmentation. The structure is shown in Fig. 3. During the training period, we train the network of attribute classification based on the attribute listings. The convolutional layers of classification network are drawn from the layers of CNN. After the attribute listings are obtained, we train the model of attribute segmentation. For each attribute, the corresponding network for segmentation is optimized based on the annotation by hand. Since the structure of CNN use the pooling layer which drops the spatial information

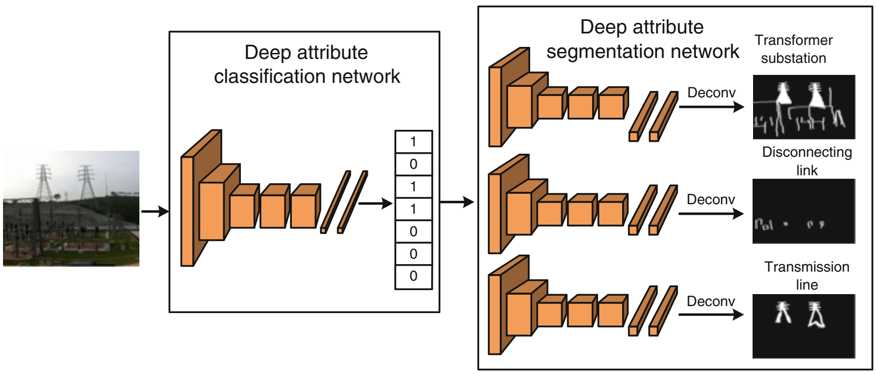


Fig. 3 The structure of deep convolutional network for attribute segmentation

of the pixel, we propose a simple and efficient network with the deconvolutional layer to recover the spatial information. To avoid the overfitting, we transfer the parameters of the convolutional layers trained on the large-scaled dataset to our model for attribute segmentation. The backward propagation and gradient descent optimization are employed to perform the optimization of the network.

Given a testing image, we first resize it to 224×224 pixels. Then, the image is normalized based on the mean value of the training images. The attribute listing is obtained by forwarding the network of classification. The corresponding networks for attribute segmentation are chosen to output the score map of the attribute.

4 Experiment

This section introduces the implementation details and the experimental setup. We also analyze the proposed network and show the results of attribute segmentation.

4.1 Implementation Details

Configuration of Network There are two parts in our network: attribute classification and attribute segmentation. The configuration of classification network is same with CNN. For brevity, we focus on introducing the configuration of the network for segmentation, which is shown in Fig. 1. Here, input indicates the input image which is resized to 224×224 pixels. There are total 6 convolutional layers in this network. The last convolutional layer outputs the attribute conditional probability map.

Training Data generation Our experiment is evaluated in the proposed attribute dataset. For each image, we provide the attribute listings and the hand annotated map of each attribute. To construct the training data for attribute classification, we randomly generate 224×224 crops from the training images. To augment the training data, the mirror version of the training image is also added. For the network of attribute segmentation, we just use the original images without crops and mirrored version since the annotation of segmentation is not changed according to these transformations. The hand-annotated map is resized to 224×224 pixels to be same with the size of the output of segmentation network.

Optimization of Network Our network is based on the Caffe framework [9]. The backward propagation is employed to compute the gradient, followed by the stochastic gradient descent optimization to update the parameter. The initial learning rate and momentum is set as 0.000001 and 0.0005 respectively. For the classification network, we initial the weights adopted from CNN which is trained on ILSVRC dataset. Then, we replace the fully connected layer fc8 with the layer, which has 7 outputs. For segmentation network, we employ the deconvolutional layer and adopt the convolutional layers of CNN.

4.2 Results of Experiments

We evaluate the proposed network on the dataset for attribute segmentation, which involves 7 attributes. We measures the scores of segmentation based on Intersection over Union (IoU) between the predicted map and the hand-annotated map.

The result is shown in Table 2. We show several qualitative results in Fig. 4. Each row shows the segmentation results of the corresponding attribute. It can be seen that our model obtains impressive performance.

Table 2 Attribute segmentation

Attribute	IoU
Transformer substation	59.15
Transformer	64.74
Disconnecting link	21.32
Transmission line	64.10
Insulator	20.09
Business hall	98.12
Person	38.31
Mean	52.26

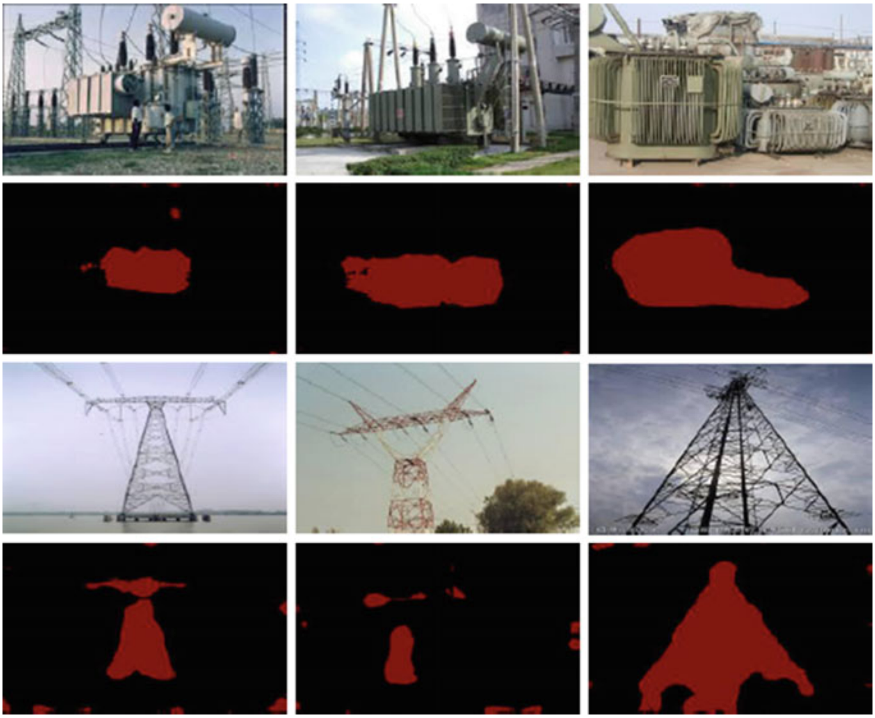


Fig. 4 The results of attribute segmentation

5 Conclusion

In this paper, we focus on attribute segmentation which outputs multi-label for one pixel. The network is composed of two parts: the convolution network for attribute classification and the deconvolution network for attribute segmentation. The result shows the advantage of our model. In future, we are interested in employing the shape prior and the relationship of attributes to obtain better results.

Acknowledgements This work was supported in part by the Natural Science Foundation of Jiangsu Province (Grants No. BK20130107).

References

1. Krizhevsky, I.S.A., Hinton, G.: Imagenet classification with deep convolutional neural networks. In: NIPS (2012)
2. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: CVPR (2014)
3. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: Decaf: a deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531 (2013)
4. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. Eprint Arxiv (2014)
5. Mostajabi, M., Yadollahpour, P., Shakhnarovich, G.: Feedforward semantic segmentation with zoom-out features. arXiv preprint arXiv:1412.0774 (2014)
6. Dai, J., He, K., Sun, J.: Convolutional feature masking for joint object and stuff segmentation. arXiv preprint arXiv:1412.1283 (2014)
7. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. arXiv preprint arXiv:1411.4038 (2014)
8. Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. arXiv preprint arXiv:1505.04366 (2015)
9. Jia, Y.: Caffe: an open source convolutional architecture for fast feature embedding. In: http://caffe.berkeleyvision.org/ (2013)

The Space Coverage Analysis Based on Planar Dynamic Grids

Ruihua Geng

Abstract In order to improve the calculating and analyzing efficiency of a satellite constellation coverage in given time, the article proposes a method which is based on planar dynamic grids to analyze the coverage of a satellite constellation. The performance of different payloads is analyzed firstly to get the restricted limit of space coverage, which provides a basis for constructing the boundary model. Secondly, the mapped relation of three-dimensional coverage to planar grid is established on the basis of coordinates projection transform, and the calculation model of space coverage criteria is brought forward. Finally, the simulation experiments validate that the models and algorithms in this article are efficacious and feasible.

Keywords Information technology · Satellite constellation configuration · Space coverage · Grid simulation

1 Introduction

With the development of space science, more and more satellites are used not only in earth observation but also in space observation. Space-based sensors can be a useful supplement to ground-based systems which are limited by atmosphere and geographical conditions. Major spacefaring nations have taken significant efforts to research and experiment on such capabilities such as the USA's MSX, SBSS, SBIRS, ODSI and Canada's MOST [1]. Traditional analyzing methods of ground coverage cannot meet the demand of space coverage analysis of onboard payloads.

Currently, major methods to calculate the coverage of satellite constellation include analytical method and grid-point coverage simulation approach (grid simulation) [2–4]. Analytical method is usually used to analyze the ground

R. Geng (✉)
Tsinghua University, Hai Dian, Beijing 100084, China
e-mail: Gengrh@sina.cn

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_64

coverage of a single satellite. However, it is difficult to calculate satellite constellation coverage since many factors such as coverage overlapping, non-circular orbit and earth rotation are not taken into account. Grid-point simulation approach is an efficient way to analyze the coverage of a satellite constellation. It is proposed by Morrison in 1973 when he was studying the coverage of a constellation including satellites in circular and elliptical orbit. The core of this approach is the sampling of grid. Usually, the earth surface is divided into many grids with the same latitude and longitude. According to simulation step size, observation information of satellites at each time point is calculated. This approach also considers the factors of different orbits and perturbation, thus statistical data of ground coverage is accurately obtained.

When the concerned object extends from 2-dimensional ground coverage analysis to 3-dimensional ground coverage, the amount of the calculation of the numerical stimulation method based on the grid sampling increases rapidly, therefore simpler algorithm needs to be considered to enhance the calculation efficiency. The literature [2] proposed several conclusions including the simply computational formula with the minimum coverage repeat cycle and the longitude and latitude coverage equivalence, i.e. the altitude is given, the latitudes of the northern hemisphere at a line of longitude are only sampled, and the essential simplicity of the periodical statistics of the coverage performance are realized based on the results, but it is noted that, the orbits of the constellation should be circular ones with the same altitude and inclination. The literature [6] used the analytical method to solve coverage time set of the single satellite, and then optimized the time-dimensional sample through dividing observable time period to enhance the speed of the calculation, and improved the performance of the coverage statistics in a small local area, however, the computational burden could be increased if used for the global coverage application.

In order to improve the calculating and analyzing efficiency of a satellite constellation coverage in given time, the article proposes a method which is based on planar dynamic grids to analyze the coverage of a satellite constellation.

2 Coverage Model of Space-Based Payloads

Different surveillance payload has different influence factors, thus different influence indices should be considered as shown in Table 1.

Using celestial equator coordinates, the origin is the earth center, and the main ring is meridian. Spatial location is represented by earth radius, right ascension and declination. The space coverage of observation is segmented through different radius settings.

The three-dimensional ball is analyzed along the center section as shown in Fig. 1.

In Fig. 1, O_e and O_s represent the earth center and spacecraft centroid respectively. The full-line circle represents free-air correction which the radius is

Table 1 Influence factors for different observation payload

Payload type	Influence factors on observation	Main influence indices
Visible light	Field of view, light, topographic barrier, earth shadow [7, 8]	The sun reflected visibility, electromagnetic wave reflection, minimum observation distance, maximum observation distance, steering angle (overlook or look-up)
Infrared	Field of view, earth's heat radiation, light, topographic barrier	Invisible heat source or deep-space background [9] (electromagnetic wave reflection), minimum observation distance, maximum observation distance, steering angle (overlook or look-up)
Radar	Transmitting power, minimum transmit-receive interval, topographic barrier	Electromagnetic wave reflection, minimum observation distance, maximum observation distance, steering angle (overlook or look-up)
Radio	Receiving sensitivity, topographic barrier, frequency band	Electromagnetic wave reflection, minimum observation distance, maximum observation distance, steering angle (overlook or look-up)

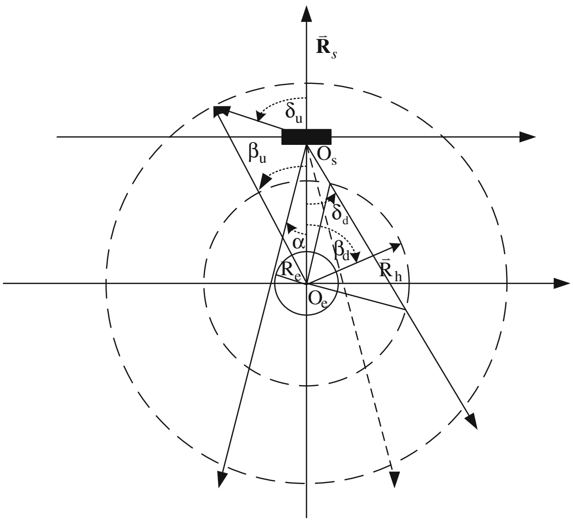


Fig. 1 Space coverage figure

represented by R_e. The dotted-line circle represents the spatial sphere which is segmented by different radius. $\overrightarrow{\mathbf{R}}_s$ is the position vector of spacecraft in geocentric equator coordinates. From O_e to O_s, δ_d representing the overlooking steering angle is defined as the angle between observation point and $-\overrightarrow{\mathbf{R}}_s$. δ_u represents looking-up steering angle is defined as the angle between observation point and $\overrightarrow{\mathbf{R}}_s$. α is the critical observation angle against deep-space background or the electromagnetic wave travels in straight lines influenced by earth shadow. Different payload has different critical observation angle. β represents geocentric angle which is defined as the angle between position vector $\overrightarrow{\mathbf{R}}_h$ and $\overrightarrow{\mathbf{R}}_s$.

In Table 1, $L_{\max}$ indicates the maximum observation distance of payloads while $L_{\min}$ the minimum observation distance. The geocentric angle value varies with different influence indices. The influence indices of radio payloads include electromagnetic wave reflection, maximum observation distance and steering angle (overlooking or looking-up). Its geocentric angle rang is as follows:

Influenced by minimum observation distance:

$$\boldsymbol{\beta}_{L\min} = \left[\arccos\left(\frac{R_s^2 + R_h^2 - L_{\min}^2}{2R_sR_h}\right), \pi\right], \quad L_{\min} \geq |R_s - R_h| \tag{1}$$

Influenced by maximum observation distance:

$$\boldsymbol{\beta}_{L\max} = \left[0, \arccos\left(\frac{R_s^2 + R_h^2 - L_{\max}^2}{2R_sR_h}\right)\right], \quad L_{\max} \leq R_s + R_h \tag{2}$$

Influenced by invisible light against deep-space background:

$$\boldsymbol{\beta}_{\alpha\min} = \left[\arccos\left(\frac{R_s^2 + R_h^2 - L_{\alpha\min}^2}{2R_sR_h}\right), \pi\right] \tag{3}$$

Influenced by electromagnetic wave reflection:

$$\boldsymbol{\beta}_{\alpha\max} = \left[0, \arccos\left(\frac{R_s^2 + R_h^2 - L_{\alpha\max}^2}{2R_sR_h}\right)\right] \tag{4}$$

In this formula, $L_{\alpha\min} = \sqrt{R_s^2 - R_e^2} - \sqrt{R_h^2 - R_e^2}, \quad R_h \leq R_s$, $L_{\alpha\max} = \sqrt{R_s^2 - R_e^2} + \sqrt{R_h^2 - R_e^2}$

Influenced by overlooking steering angle:

$$\boldsymbol{\beta}_{\delta_d} = \begin{cases} [0, \pi], & R_e \leq R_h < R_s \sin\delta_d \\ \left[0, arc\cos\left(\frac{R_s^2 + R_h^2 - L_{\delta_d\min}^2}{2R_sR_h}\right)\right] \cup \left[arc\cos\left(\frac{R_s^2 + R_h^2 - L_{\delta_d\max}^2}{2R_sR_h}\right), \pi\right], & R_s \sin\delta_d \leq R_h < R_s \\ \left[arc\cos\left(\frac{R_s^2 + R_h^2 - L_{\delta_d\max}^2}{2R_sR_h}\right), \pi\right], & R_s \leq R_h \end{cases} \tag{5}$$

In this formula, $L_{\delta_{d\min}} = \sqrt{R_h^2 - (R_s \sin\delta_d)^2} - R_s \cos\delta_d$, $L_{\delta_{d\max}} = \sqrt{R_h^2 - (R_s \sin\delta_d)^2} + R_s \cos\delta_d$ ∘.

As for optical payloads, front light observation should also be considered as a influence index, i.e. the angle between optical axis and solar ray is less than 90°.

3 The Space Coverage Analysis Based on Planar Dynamic Grids

Currently major methods to calculate the coverage of satellite constellation include analytical method and grid-point coverage simulation approach. It is difficult to calculate satellite constellation coverage using analytical method. Grid-point simulation approach is also limited by discrete sampling while accurate calculation requires intensive sampling. The space coverage analysis based on planar dynamic grids makes full use of the advantages of analytical and grid-point coverage simulation approach. It simplifies the calculation and meets the demand of coverage analysis of complicate satellite constellation.

As for the coverage model of a single satellite, the angle between the satellite $\overrightarrow{\mathbf{R}}_s$ and any point on a R_h radius virtual sphere is

$$\cos\beta_p = \frac{\overrightarrow{\mathbf{R}}_p \cdot \overrightarrow{\mathbf{R}}_s}{\left\|\overrightarrow{\mathbf{R}}_p\right\|\left\|\overrightarrow{\mathbf{R}}_s\right\|} = \cos\varphi_p \cos\varphi_s \cos(\lambda_p - \lambda_s) + \sin\varphi_p \sin\varphi_s \tag{6}$$

Here, φ and λ is declination and right ascension. As the geocentric angle is determined, the relationship of the point coordinate of the geocentric angle satisfied in the sphere is follows according to the above formula:

$$\lambda_p = \pm\arccos\left(\frac{\cos\beta_p - \sin\varphi_p \sin\varphi_s}{\cos\varphi_p \cos\varphi_s}\right) + \lambda_s \tag{7}$$

According to the rotation of the Earth, the spatial position represented by geographical longitude and latitude $\left[\varphi_{L,p}, \lambda_{L,p}\right]$ can be obtained only by the translation from the above right ascension coordinate relative to the angle of the rotation of the Earth.

If different influence indices are considered, $\boldsymbol{\beta}_c$ indicates the value set of geocentric angle of space coverage. Assuming $\boldsymbol{\beta}_c$ is composed of K continuous collection interval among which $\beta^L_{c,k}$ and $\beta^U_{c,k}$ indicates the upper and lower limit. According to Table 1, different type of payload has its own influence indices. Geocentric angle of each type is represented respectively as follows:

Visible Light: $\boldsymbol{\beta}_c = \boldsymbol{\beta}_{L\min} \cap \boldsymbol{\beta}_{L\max} \cap \left(\boldsymbol{\beta}_{\delta_u} \cup \left(\boldsymbol{\beta}_{\alpha\max} \cap \boldsymbol{\beta}_{\delta_d}\right)\right)$
Infrared: $\boldsymbol{\beta}_c = \boldsymbol{\beta}_{L\min} \cap \boldsymbol{\beta}_{L\max} \cap \left(\boldsymbol{\beta}_{\delta_u} \cup \left(\boldsymbol{\beta}_{\alpha\max} \cap \boldsymbol{\beta}_{\alpha\min} \cap \boldsymbol{\beta}_{\delta_d}\right)\right)$
Radar: $\boldsymbol{\beta}_c = \boldsymbol{\beta}_{L\min} \cap \boldsymbol{\beta}_{L\max} \cap \left(\boldsymbol{\beta}_{\delta_u} \cup \left(\boldsymbol{\beta}_{\alpha\max} \cap \boldsymbol{\beta}_{\delta_d}\right)\right)$
Radio: $\boldsymbol{\beta}_c = \boldsymbol{\beta}_{L\max} \cap \left(\boldsymbol{\beta}_{\delta_u} \cup \left(\boldsymbol{\beta}_{\alpha\max} \cap \boldsymbol{\beta}_{\delta_d}\right)\right)$

Of course, it can do a constraint selection according to the actual need, For instance, the pitch is not be considered at certain application, and therefore the $\boldsymbol{\beta}_{\delta_u}$ constraint in above equation can be eliminated. Such set operation is also suitable for other synthetic analysis of the constraint in different orbital planes. However, for the common constraint between different orbital planes, it can't be described by the simple set operation due to the different reference points of the geocentric angle, therefore it should be expressed by the method based on the 2-dimensional dynamic grid.

First, the isometric projection of the geographical longitude and latitude is made for the celestial sphere with a radius R_h at the time t, and then mapped for 2-dimensional plane. The horizontal axis stands for the longitude with a mapping range [−180°, 180°] and the vertical axis stands for the latitude with a mapping range [−90°, 90°]. The plane is gridded and expressed as the matrix $I_{t,h}(m,n)$ with $M \times N$ dimensions. For the convenience of the expression and analysis of the matrix, M and N are chosen as even, and therefore the longitude and latitude mapped by the every unit (m, n) is $(\varphi_L, \lambda_L) = \left(\left(\frac{M+1}{2} - m\right)\Delta\varphi_L, \left(n - \frac{N+1}{2}\right)\Delta\lambda_L\right)$.

The matrix $I_{t,h}$ is extended to the matrix $I_{t,h}^{sat}$ and $I_{t,h}^{sun}$ that is used for the constraint analysis of the coverage of the satellite and the Sun respectively. First, the matrix $I_{t,h}^{sun}$ is given assignment, and then the unit section of the vertical axis of the matrix is designated according to the upper limit $\beta_{c,k}^{U}$ that covers the continuous set section of the geocentric angle.

$$\mathbf{m}^U = \left\{ m \middle| \max\left(-\frac{\pi}{2}, \varphi_s - \beta_{c,k}^U\right) \leq \left(\frac{M+1}{2} - m\right)\Delta\varphi_L \leq \min\left(\varphi_s + \beta_{c,k}^U, \frac{\pi}{2}\right) \right\} \quad (8)$$

For $m \in \mathbf{m}^L$, the φ_L is obtained, and further the two solutions $\lambda_{L,-}$ and $\lambda_{L,+}$ of λ_L can be calculated and changed into the expression of the unit abscissa $n_+^U(m) = \max_n\left\{\left(n - \frac{N+1}{2}\right)\Delta\lambda_L \leq |\lambda_{L,+} + \pi|_{2\pi} - \pi\right\} \& \, n_-^U(m) = \min_n\left\{\left(n - \frac{N+1}{2}\right)\Delta\lambda_L \geq |\lambda_{L,-} + \pi|_{2\pi} - \pi\right\}$. Here, $|\bullet|_{2\pi}$ stands for modular arithmetic, and the unit section of horizontal axis of the matrix is follows:

$$\mathbf{n}^U = \begin{cases} \left\{ n \middle| n_-^U(m) \leq n \leq n_+^U(m), m \in \mathbf{m}^U \right\}, & n_-^U(m) \leq n_+^U(m) \\ \left\{ n \middle| n_-^U(m) \leq n \leq N, m \in \mathbf{m}^U \right\} \cup \left\{ n \middle| 1 \leq n \leq n_+^U(m), m \in \mathbf{m}^U \right\}, & n_-^U(m) > n_+^U(m) \end{cases} \quad (9)$$

In a similar way, the unit section of the vertical axis of the matrix $\mathbf{m}^L$ can be obtained according to the lower limit $\beta_{c,k}^{L}$ that covers the continuous set section of the geocentric angle. For $m \in \mathbf{m}^L$, the according unit abscissa section is expressed as $\mathbf{n}^L$. From the above analysis, the assignment of the matrix $I_{t,h}^{sat}$ is follows:

$$\mathbf{I}_{t,h}^{sat}(m,n) = \begin{cases} 1, & \left\{(m,n) \middle| n \in \mathbf{n}^U, m \in \mathbf{m}^U\right\} - \left\{(m,n) \middle| n \in \mathbf{n}^L, m \in \mathbf{m}^L\right\} \\ 0, & else \end{cases} \quad (10)$$

In a similar way, the coverage constraint of the Sun $I_{t,h}^{sun}(m,n)$ can be obtained, and the coverage constraint of the single satellite $I_{t,h}$ is expressed as

$$\mathbf{I}_{t,h}(m,n) = \mathbf{I}_{t,h}^{sat}(m,n) \times \exp\left[\gamma \ln \mathbf{I}_{t,h}^{sun}(m,n)\right], m \in [1,M], n \in [1,N] \tag{11}$$

Here, $\gamma = 0$ stands for the visible constraint that is free of the solar reflection, and $\gamma = 1$ stands for the visible constraint that is not free of the solar reflection. From the above constraint coverage model of the single satellite, we can obtain the constraint coverage of the constellation. Given the constellation is comprised of satellites or spacecrafts of number K, and therefore the 2-dimensional grid of the constellation coverage $I_{t,h}$ is expressed as

$$\mathbf{I}_{t,h}(m,n) = \sum_{k=1}^{K} \mathbf{I}_{t,h}^{sat_k}(m,n) \times \exp\left[\gamma_k \ln \mathbf{I}_{t,h}^{sun}(m,n)\right], m \in [1,M], n \in [1,N] \tag{12}$$

4 Indices of Constellation Space Coverage

Be confined to the article length, according to the different applications of the constellation, this article only defines the index of the k multiplicity coverage of constellation relative to space to evaluate the work performance of the constellation, and meanwhile makes a verification for the analytic method of the space coverage based on the 2-dimensional dynamic grid model. According to the surface integral, the area of the different sampling unit in the same layer gird $S(m,n)$ is:

$$s(m,n) = \Delta\lambda \cdot \left\{ \sin\left[\varphi_L(m,n) + \frac{\Delta\varphi}{2}\right] - \sin\left[\varphi_L(m,n) - \frac{\Delta\varphi}{2}\right] \right\} \tag{13}$$

Assume the grid point set in the interested zone as **P**, and therefore the instantaneous k multiplicity coverage is:

$$C_{t,h,k} = \frac{\sum_{(m,n)\in\mathbf{P}} b_k(m,n) \cdot s(m,n)}{\sum_{(m,n)\in\mathbf{P}} s(m,n)}, \quad b_k(m,n) = \begin{cases} 1, & \mathbf{I}_{t,h}(m,n) \geq k \\ 0, & else \end{cases} \tag{14}$$

5 Simulation Experiment

Space-based Surveillance System (SBSS) is a LEO optical observation constellation comprised of 4-8 satellites. At present, a SBSS-1 satellite have been launched, its payload is a visible light sensor with high sensitivity that installed in the biaxial universal gymbal with a high speed of rotation [10], have the ability of good observation for GEO and short repeat cycle and all-weather, etc. We take the characteristics of the orbit of SBSS-1 as the simulation reference, the parameters are set as: the altitude is 630 km, the inclination is 98 degree, and the eccentricity is 0.001. Considering the deployment and backup of the constellation, we choose four satellites distributed uniformly along the different right ascensions of ascending node to compose the constellation and make the performance simulation of the space coverage. The start and stop time of the simulation is 2014-9-1 12:00 UTC and 2014-9-2 12:00 UTC respectively, the step is 60 s, the payload is the visible light vehicle, the steering angle of the pitch is 90°, the shortest detection range is 0 km, the farthest detection range is 50,000 km, and the size of dynamic grid is 180 × 360.

5.1 STK Simulation Demonstration

Satellite Tool Kit (STK) is the top of the analytic software for aerospace, and it can be chosen as the comparison and verification of simulation for the method used in this article. We take the first sample epoch (1 Sep 2014 12:00:00.0) and the Earth's surface to make the coverage projection of the single satellite and the sunlight. The results calculated by the method in this article are shown in Figs. 2a and 3a, and the results calculated by STK are shown in Figs. 2b and 3b.

From the comparison for the position and size of the coverage zone of the single satellite and the sunlight in Figs. 2 and 3, we can see the results calculated by the analytic method of coverage zone based on the model used in this article are the same as the ones by STK, and the coverage range of the longitude and latitude at the same epoch are the same. It is shown that the method proposed by this article is correct and effective, and it is not only used for the analysis of the ground coverage performance, but also extended to the space of different altitudes to make the coverage simulation and calculation and to realize the coverage analysis of all the space.

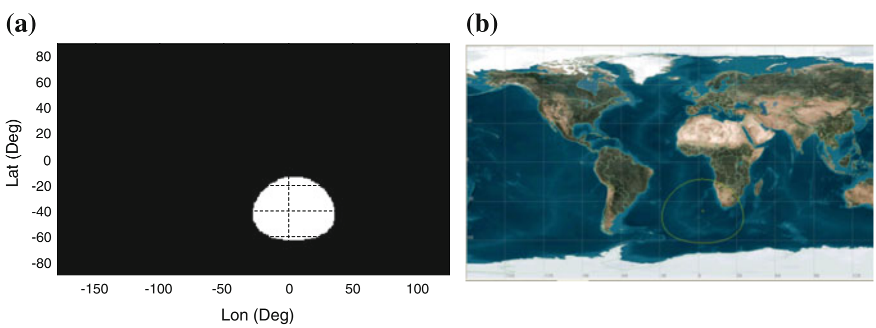


Fig. 2 Results of the coverage of the single satellite. **a** calculated by the method in this article. **b** The results calculated by STK

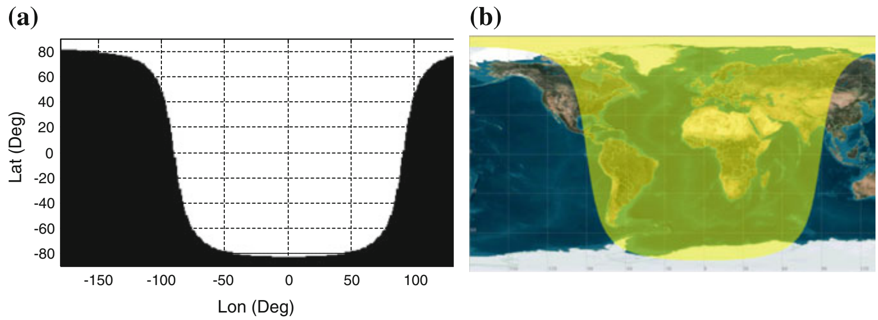


Fig. 3 Results of the coverage of the sunlight. **a** Calculated by the method in this article. **b** Calculated by STK

5.2 *Surveillance Ability Simulation*

The GEO is the typical orbit of space object, and also the primary observation object of space-based optical constellation. We take the altitude zone of 42165 km (GEO) from the center of the Earth to make the coverage analysis. Be confined to the article length, we only give the coverage figures at the first sample epoch and the middle sample epoch in one day, as shown in Fig. 4. In the figure, the black zone indicates not covered, and the white zone indicates covered.

Due to the inclination, the simulation time, the solar position and the frontlighting observation, it presents the asymmetric shape at the north and south in the coverage zone. The thin strips emerged on the edge and the asymmetric zone at left and right is because of the frontlighting. The black circular zone are the part of the shadow, and can‘t be visible. In view of the simulation of space-based optical constellation relative to the space coverage of the synchronous belt, the constellation has the adequate ability to cover the synchronous belt completely, however, the specific observation range is related to the payload and the observation strategy, etc.

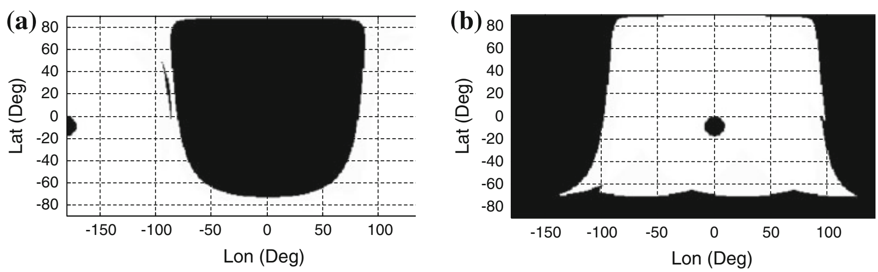


Fig. 4 Gird coverage figure of constellation of 42165 km altitude. **a** 1 Sep 2014 12:00:00.0. **b** 1 Sep 2014 24:00:00.0

5.3 The Space Coverage Rate of Different Altitudes

For the convenience of analysis, we take the sample in segments for the space of different altitudes. The sampling interval of altitude from 6400 to 7400 km is 10 km, and the interval of altitude from 7400 to 47400 km is 1000 km. During the setting of the constellation, we choose four satellites to make the simulation, therefore one region can be observable by four satellites at the same time in the most (coverage multiplicity k = 4). Here, we calculate the instantaneous space coverage rate (first sampling epoch) and average space coverage rate of different altitude ranges and coverage multiplicities, the results are shown in Fig. 5.

From Fig. 5, we can see the multiplicity of the highest coverage rate is 1 in any altitude, and the lowest is 4. The instantaneous space coverage of the constellation are all different at any time. From the instantaneous coverage at the first sample epoch, we can see all regions maybe not able to reach the coverage of 3 multiplicity or 4, therefore the coverage rate is zero. With the increasing altitudes, the visible region of the payload is expanding, and the coverage rate of 3 and 4 multiplicity is also increased. It is noted that the instantaneous coverage rate of 1 multiplicity goes to decline while achieving at a critical value. The primary reason of the phenomenon is the fact that considering the front lighting constraint, and the opposite black lighting region of the optical payload are all not be covered, therefore the coverage rate of payload relative to space is affected. The Averaged coverage rate makes a statistical average for all day coverage rate, it shows that the coverage rate of constellation relative to space is almost increasing gradually with the increasing altitudes of space. The statistical results further validate the effectiveness of the method and the according model proposed in this article.

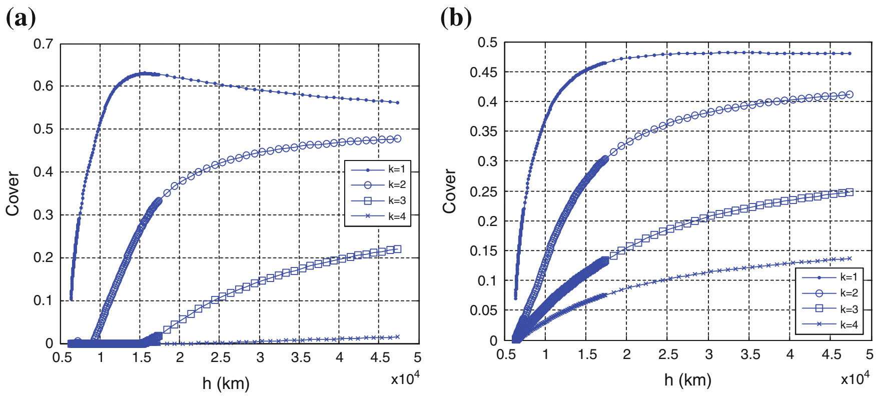


Fig. 5 Space coverage rate of different altitudes. **a** Instantaneous space coverage rate at 1 Sep 2014 12:00:00.0. **b** Average space coverage rate

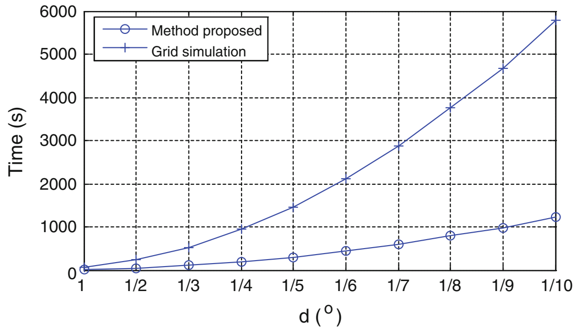


Fig. 6 The comparison of the consuming time with the point coverage numerical simulation

5.4 *Comparison with the Grid Point Numerical Simulation*

While calculating the space coverage, we increase the sampling interval of dynamic grid from d = 1° to d = 0.1°, i.e. subdivide the sampling unit from 1 grid to 10 grids. Figure 6 shows the results of the consuming time for the method proposed in this article and the method of the point coverage numerical simulation.

When the sampling interval is 1°, the unit amount of all grids is to be 6.48×10^4, and the time consuming by the method in this article and by the point numerical simulation is 10.4 s and 55.1 s respectively. When the sampling interval is 0.2 degree, the unit amount of all grids is to be 1.62×10^6, and the time consuming by the method in this article and by the point numerical simulation is 303.4 s and 1464.7 s respectively. When the sampling interval is 0.1 degree, the unit amount of all grids is to be 6.48×10^6, and the time consuming by the method in this article and by the point numerical simulation is 1216.7 s and 5775.6 s respectively. It is shown that the calculation speed of the method in this article is five times faster than the point numerical simulation. Therefore, compared with the point numerical simulation, the method in this article combines the analytic method and avoids the data sample in the longitude dimension, and enhances the calculation speed effectively, especially the sampling density is much higher, the performance is improved much obviously.

6 Conclusion

Based on the analysis of the restricted condition of space-based payload, this article builds a coverage constraint model of the payload, and combines the advantages of both the analytic method and the point coverage numerical simulation to propose a method to analyze the space coverage on the basis of 2-dimensional dynamic grid and validate its correctness. Besides, this article proves that the method can effectively realize the performance analysis of the space coverage of the constellation through the numerical simulation and comparison of the space-based constellation space coverage, and can enhance the calculation speed obviously. Therefore, the method can be applied for the coverage analysis and capability assessment of the SBSS conveniently.

References

1. Wang, J.J., Yu, X.H.: The current development and characteristics analysis of foreign space-based space surveillance. J. Acad. Equip. Command Technol. **17**(4), 33–37 (2006)
2. Han, C., Deng, L., Xu, J.: Improved grid simulation for constellation coverage performance evaluation. Comput. Simul. **12**(22), 21–23 (2005)
3. Jian, P., Zou, P., Xiong, W., Chen, Z.: Improved grid method for analysis on coverage performance of staring sensors based LEO. J. Air Force Eng. Univ. (Nat. Sci. Ed.) **13**(3), 35–39 (2012)
4. Wang, B.,An, W.,Zhou, Y.-Y.: Analysis on airspace coverage performance of tracking sensors. Space Control **27**(6), 90–95 (2009)
5. Gao, H.-M., Xu, X.-H.: Design of regional coverage satellite constellation based on analytical method. J. Mod. Def. Technol. **40**(2), 24–26 (2012)
6. Zhang, Q., Zhao,Y., Xu, M.: Computation model of constellation space coverage performance. J. Spacecr. TT&C Technol. **30**(1), 6–10 (2011)
7. Liu, X.C., Liao, Y., Wen, Y.-L.: Applying the improved genetic algorithm to optimize the orbit for space-based optics surveillance platform. J. Natl. Univ. Def. Technol. **36**(3), 58–62 (2014)
8. Liu, X.C., Liao, Y.L., et al.: Simulation and analysis of space object visibility based on space-based optics surveillance. In: Proceedings of International Conference ICSC2012, pp. 195–203 (2012)
9. Deng, Y., Wang, C.M., Zhang, Z.-Z.: Analysis on coverage performance of staring sensors infrared LEO constellation. J. Astronaut. **32**(1), 123–128 (2011)
10. Cui, X.-X.: A survey of space-based surveillance of USA. Space Int. **7**, 37–43 (2011)

Multi-sentence Level Natural Language Generation for Dialogue System

Junkuo Cao, Guolian Chen, Lihua Wu, Yu Zhang and Ziqiang Luo

Abstract In multi-sentence level natural language generation (NLG) system, the first task is to classify the constraint-value pairs into several groups. Following that, each group will be translate to a talk session. In this paper, we propose three classification algorithms. The first method looks for talk session group directly, the second one applies sentence frame search, and the third one utilizes dynamic programming. After classification, we propose sentence level and phrase level generation method. Although these algorithms are not sophisticate, they work well due to easy training, fast response and high quality. Dialogue system for operating in-car devices and services require fast system response time, and some PND applications are rare in memory/computational resources, so this method would be a good choice for these environments.

Keywords Natural language generation · Dialogue system · Dynamic programming

1 Introduction

The mainstream characterization of NLG addresses the problem of mapping non-linguistic representations to expressions in natural language [1]. NLG researchers have dealt with generation from a variety of non-linguistic representations including database entries, formulas in logical calculi and expressions of knowledge representation formalism. Generally, the task of generating natural language can be broadly separated into two sub tasks: (1) determining the content ("what to say") and (2) determining its realization ("how to say") [2]. Generation of linguistic expressions of various sizes and types has been investigated, including substantial units (e.g., noun phrases), individual sentences and multi-sentence

J. Cao · G. Chen (✉) · L. Wu · Y. Zhang · Z. Luo
Department of Computer Science and Technology, Hainan Normal University, Haikou 571158, China
e-mail: 276565829@qq.com

B.-Y. Cao et al. (eds.), *International Conference on Oriental Thinking and Fuzzy Logic*, Advances in Intelligent Systems and Computing 443,
DOI 10.1007/978-3-319-30874-6_65

discourse. Michael Elhadad presented a general approach to lexical choice that can handle multiple, interacting constraints, which focuses on the problem of floating constraints, semantic or pragmatic constraints [3, 4]. The sentence planning stage is embodied by the SPoT sentence planner [5] while the surface realization stage is embodied by the FERGUS surface realizer [6, 7]. John Chen extended their works, which discuss a NLG system that is composed of SPoT and FERGUS [8].

In this paper, we apply n-gram model, which has the considerable advantages of lending robustness to NLG systems, and making automatic adaptation to new domains with poor resource, to realize multi-sentence discourse NLG.

We now present our NLG System in the following sections. In Sect. 2, we introduce our training database. Following that, in Sect. 3, we will elaborate the three classification algorithms for constrains queue. Then in Sects. 4 and 5, we describe sentence level generation and phrase level generation respectively. Finally, we give our conclusion in Sect. 5.

2 Training Database

In order to achieve multi-sentence level NLG, we labeled a corpus which comes from real dialogue of restaurant query between customers and telephonist. Every sentence in the corpus has been annotated as semantic representation which includes two parts: (1) speech action, and (2) semantic slots information. The first one is used to define the sentence pattern, and the second shows the content. This tedious and heavy labeling work could be done by most workable dialog systems, which have the high accuracy of the constraint identifier. In the training process, each sentence will be mapped into constraint-value pairs sequence. And the talk session, in the same time, can also be mapped into these pairs sequence by only insert a splitting tag between the adjacent sentences. Let's denote constraint as c, sentence as s and session as S. So each session S could be wrote as $S = \{s_1\#s_2\#\cdots\#s_n\}$, and sentence s could be as $s = \{c_1||c_2||\cdots||c_m\}$. Apparently, the tag "#" is the sentence conjunctive symbol, and the tag "||" is constraint splitting character. In the training process, we can easily get $P(c)$, $P_{start}(c)$, $P_{end}(c)$ and $P(c_i|c_j)$ by n-gram model. $P(c)$ and $P(c_i|c_j)$ are clear literally. $P_{start}(c)$ and $P_{end}(c)$ represent the probability of constraint c being as the star or end of one sentence. Meanwhile we can also get $P(s)$, $P_{start}(s)$, $P_{end}(s)$ and $P(s_i|s_j)$. In the statistic process, we also construct database for NLG system, which includes three tables. They are respectively table constraint, lexical and sentence. The table constraint represents all constraint-value pairs in the corpus. Table lexical stored all the lexical information for the each constraint-pair. And table sentence records all natural language sentences information for each semantic representation, which include SpeechAction and the sequence of constraint-value pairs.

3 Classification Model

The first task of multi-sentence discourse NLG is to classify the input non-linguistic representation, the sequence of constraint-value pairs concerned about subject, into several groups. Each group represents a natural sentence. In order to achieve this task, three classification algorithms have been used. They are Session Search (NLG_C1), Sentence Search (NLG_C2) and Constraint Search (NLG_C3). In classification process, all classification models just consider constraint name and order, while constraint value and its' lexical expression have been ignored. These constraints involved 12 types which list as shown Table 1. According to real dialogue data analysis, we assume that all constraint type appeared only once in constraints queue, so we can translate the constraint sequence into binary sequence by the specific position of the order constraint, and then corresponds to a decimal number. For example, constraint queue {dresslevel: casual ||pricelevel: cheap|| cousine: Chinese} could be transformed 110000000001, and the corresponding decimal will be 3073.

We use session search first. If a session matches nothing from the corpus, sentence search will be used after then. Generally, session search and sentence search could work well if the corpus have exactly the same session or have some sentence frames which can make up the whole session. But most time we should use constraint search to construct a session by dynamic programming algorithm.

Table 1 Contraint type and example

Number	Constraint type	Example
1	restaurant:hasCuisine	Chinese restaurant
2	restaurant:hasPriceLevle	Cheap
3	restaurant:hasName	Barefoot cafe bar
4	restaurant:hasService	Excellent
5	restaurant: hasBusinessHours	Lunch
6	poi:hasCity	Lincoln
7	poi:hasStreetName	Ross road
8	restaurant:hasRating	Five star restaurant
9	restaurant:hasSpecialCon	Spicy food
10	restaurant: accpetsCreditCards	Yes
11	restaurant:hasReservation	Yes
12	restaurant:hasDressCode	Casual

3.1 NLG_C1

In our corpus, all session will be stored with a decimal identity number and 12 binary sequence. Given a semantic constraint representation $C = \{c_1 \| c_2 \| \cdots \| c_m\}$, we translate the sequence into binary queue, and then decimal number. NLG_C1, the session search algorithm, just count *Count*(*S*), while *Count*(*S*) is the occurred number of the session *S* which has the same decimal identity with the input constraint queue *C*.

3.2 NLG_C2

NLG C2 is based on sentence search, rather than on session search used in NLG_C1. In this algorithm, we also translate the input constraint queue *C* into binary sequence B_C. And then we search all possible sentences set $S = \{s_1, s_2, \ldots, s_k\}$, in which each s_i match part of the constraint *C*. Finally according to NLG_C2 algorithm to pick up a subset S' from *S* based on the following two conditions:

(1) each binary sequence from S′ should partially match B_c
(2) the whole binary sequence made by S′ should fully match B_c (Fig. 1)

3.3 NLG_C3

In NLG_C3 algorithm, we construct a tree by the given constraint sequence $C = \{c_1 \| c_2 \| \cdots \| c_m\}$. Each node in the tree is one of constraint within *C*. Generally, a talk session representation *C* includes about 2–8 constraint. In the tree construction

Input: $C=\{c_1, c_2, \cdots c_n\}$
Output: $B_{s1}, B_{s2} \cdots B_{sm}$ (B_{sm} is binary sequence for sentence s_m)
Algorithm:
1. Translate *C* into binary sequence B_c
2. Find out the all possible sentence set $S=\{s \mid (B_s \text{ xor } B_c) \text{ and } B_s=0\}$
3. Find out all possible $S' \subseteq S$, $\cup B_{si}=B_c$, $\cap B_{si}=0$ and $s_i \in S'$
4. Calculate the weight for each S’, $\text{Weight}(S')=\sum_{s_i \in S'} count(s_i)$,
   while count(s_i) is occurred times of sentence s_i in the corpus.
5. Find the *S’* which has largest Weight, and return $B_{s1}, B_{s2} \cdots B_{sm}$, while $s_i \in S'$

Fig. 1 The algorithm of NLG_C2, sentence search

1. Initial a virtual node as tree root, set it's height as -1 and trace as an empty string, and set it's search cost as 1.0.
2. rank all constraints c by p_{start}(c), and put top N node into need extend queue Q.
3. While Q is not empty
    3.1 remove the first node c from Q
    3.2 choose N best node as c's children, the first N/2 children as sentence continual part, while the remainder as another sentence start part.
    3.3 set the N best node height as c's height pluses 1, and set their trace as c's trace appends splitting tag and children constraint.
    3.4 calculating the N best node search cost.
    3.5 add the N best node into extend queue
    3.6 go to step 3.

Fig. 2 The algorithm of NLG_C3, constraint search

process, we add a virtual node as root. Considering the computing complexity, we just select N-Best node as children, instead of constructing the complete tree. Given current node c, we would choose N nodes c_c as its children. The first N/2 nodes depend on P(c_c|c), and the last N/2 nodes depend on $P_{start}(c_c)$. The algorithm of children node selection can be shown in Fig. 2. While children node search cost are calculated as the following formula:

$$\begin{array}{ll} c_c.\text{cost} = c.\text{cost} \times P(c_c|c) & \text{if } c_c \text{ and } c \text{ in same sentence} \\ c_c.\text{cost} = c.\text{cost} \times P_{\text{start}}(c_c) & \text{if } c_c \text{ is another sentence start} \end{array}$$

After the tree construction, we choose the maximum search cost valid leaf node and let its' trace as the final output.

3.4 Experiment on Classification

The purpose of the paper addresses to generate acceptable sentence frame by semantic constraint pairs. To realize the purpose, we labeled the real dialogue of restaurant query between users and telephonist. In the all annotation corpus, including 19 users, there are 231 sessions which totally include 848 sentences. According to the above constraints, each session and sentence could map into a constraint sequence. And on the other hand, we can also map constraint set into a session or a sentence directly.

To evaluate the defectiveness of our NLG system, we design an experiment on different size corpus. The first test uses the first ten dialog session as training data, and the last nine sessions as testing data. In the second test, we increase the size of the training data, in which the first 16 dialog sessions will be the training data, and the remaining sessions will be testing data. As shown in Tables 1 and 2, although the algorithm NLG1 gets the highest precision, the recall is lowest. While the model

Table 2 The first 10 users dialog as training data, the last 9 ones as testing data

Algorithm	Recall	Precision
NLG_C1	16.2	74.9
NLG_C2	24.3	58.3
NLG_C3	31.8	39.6
NLG_C1 + NLG_C2	24.3	62.7
NLG_C1 + NLG_C3	31.8	52.9
NLG_C1 + NLG_C2 + NLG_C3	34.7	55.9

Table 3 The first 16 users dialog as training data, the last 3 ones as testing data

Algorithm	Recall	Precision
NLG_C1	21.3	79.7
NLG_C2	31.0	65.1
NLG_C3	34.4	41.3
NLG_C1 + NLG_C2	34.3	76.4
NLG_C1 + NLG_C3	37.8	56.9
NLG_C1 + NLG_C2 + NLG_C3	39.5	59.6

NLG3 gets highest recall, the precision is much worse than NLG1. So, we also try to combine these models by specific order. For example, the model NLG1 + NLG2 means that we use NLG3 after NLG1 does not work well. Apparently, the combined algorithm is effective than the single one. In the same time, the test result show that our training model can still improve effectiveness with bigger corpus (Table 3).

4 Sentence Level Generation

After semantic representation is classified, we should generalize a sentence frame for each group. This work is called sentence level generation. This time SpeechAction and constraint value have been involved. For example, given semantic representation, SpeechAction = PStatements, num = twenty, cuisine = fast food, takeCreditCard = yes, the sentence frame something like "there are [num = twenty] [cuisine = fast food] that [takeCreditCard = yes]" should be generalized.

Because of data sparse, we generalize 2^n sentences frame for each sentence $s = \{c_1 = v_1, c_2 = v_2 \ldots, c_n = v_n\}$. For example: the sentence [hasName = don felix cafe] is [distance = within two miles] could be generalized four sentence as follows:

(1) don felix caf is within two miles.
(2) [hasName] is within two miles.
(3) don felix cafe is [distance].
(4) [hasName] is [distance].

Generally the given constraint-value pairs sequence is too strict to find any sentence frame. So we also use generalization method for sentence Frame Search Algorithm that relaxes these conditions step by step. Unfortunately, there is still sometimes no sentence frame for the most relaxative constraint-value pairs. In this time, we will use a split algorithm to split the search constraint-value pairs into several small ones, and then search their appropriate sentence frame.

5 Phrase Level Generation

5.1 Off-Line Mode

In the off-line mode, all user utterances available for training are collected. Then, two different approaches can be taken: one is called parse-tree-based approach, and the other is called chunking-based approach. In the parse-tree-based approach, all the utterances from the users are structurally annotated as parse-trees. The lowest level phrase nodes that contain user constraints or potentially useful phrases for the system responses can be identified. In case of the chunking-based approach, the user utterances are tagged by a POS tagger, chunked, or even parsed with a deep structure parser. Combining the chunker results with the constraints found by the constraint identifier, the phrase nodes that contain user constraints can be extracted. For both methods, extracted phrases form a pool of original examples of how to lexicalize the constraints are collected in a database. They can be further generalized into phrase patterns, if the constraints are replaced by their semantic class names. For example, the generalized phrase pattern some CuisineType food is generated by replacing Chinese with CuisineType in the original phrase some Chinese food. The node instances in the pool can also be re-composed to generate new pattern/phrase combinations by some node operations. The combination can be guided by X-bar theory, for example. Thus, we can construct a response database which includes original phrase nodes, general phrase patterns and the results of the node operations. During the conversation, response generator will select phrase from this database that correspond to the given constraints.

5.2 On-Line Mode

In the online mode, we may use the chunking-based approach to extract the phrases the user uses in a constraint specification. Then, we may update these phrases with values retrieved from the knowledge base or database. For example:

User: I want a five star restaurant that serves Chinese food.

System: I did not find any five star restaurants that serve Chinese food, but I found 20 four star restaurants that serve Chinese food.

Note: The first phrase can be constructed by using the user input, and the second expression can be produced by generalization over the Rating constraint.

During the conversation, when there are multiple candidates, selection criteria and back-off strategies will have to be used. For example, when a combination of constraints and values has been seen in the original corpus (possibly with a count over a certain threshold), the generation module can use that pattern. When the constraint combination has been seen before, but with different values, a generalization will take place. We attempt to find patterns that overlap as much as possible with the input. For example, if the two given constraints are CuisineType:Indian and RatingLevel:5star and we have never seen this combination in our corpus, we will try to relax each constraint by employing general patterns like: a CuisineType restaurant with five star rating or an Indian restaurant with RatingLevel rating. Furthermore, general patterns which relax all the constraints will also be employed: a Cuisine-Type restaurant with RatingLevel rating, a RatingLevel restaurant that serves CuisineType food, etc. In the last case, when the constraints combination has not been seen before by itself or in a larger expression, we attempt to use patterns created by combining the existing patterns. At the end, the proposed patterns can also be ranked according to a language model (the training corpus and external sources might be used for deriving the language model), to ensure alignment with the user utterance or variability of the systems output. One interesting aspect of the problem is that the lexicalization of different constraints depends on the particular constraint and on the values for that constraint. For example, the location constraint is best realized with a PP, like a restaurant in Lincoln, while a Lincoln restaurant is bad, even though this is a good pattern for most constraints. Similarly, depending on the value of the constraint, different patterns are preferred. For example, if the CuisineType is Barbeque we need to select a realization different than the one for CuisineType = Chinese: a restaurant that serves Chinese food is good, but a restaurant that serves Barbeque food is not a very fluent form.

6 Conclusion

This paper presents a multi-sentence level language sentence generating system. In the classification process of constraint queue, we proposed three algorithms, including session search (NLG_C1), sentence search (NLG_C2) and constraint search (NLG_C3). According to search result, we try to combine these classifications methods. The experiment data show that the combined algorithm can get more reliable result. For sentence level generation, we put forward a gradually generic method to relax the constraints step by step, and then construct complete sentence structure by appropriate change based on generalization of semantic constraint. In the phrase generation, we use the online mode and offline mode respectively, which provide a viable solution for the phrase presentation.

Acknowledgments Thanks to the support by National Natural Science Foundation (No. 61363032, 61262077, 61462025, 61463012), Major Scientific Projects of Haikou (No. 2012-050). And this work is supported by Hainan International Cooperation Key Project (KJHZ2014-25); Hainan Special Project for Social Development of Science and Technology (SF201436), and Major Scientific Projects of Haikou (No. 2012-050).

References

1. Mellish, C., Scott, D., Cahill, L., Paiva, D., Evans, R., Reape, M.: A reference architecture for natural language generation systems. Nat. Lang. Eng. **12**(1), 1–34 (2006)
2. Stent, A., Bangalore, S.: Natural Language Generation in Interactive Systems. Cambridge University Press, Cambridge (2014)
3. Goldberg, Y., Tsarfaty, R., Adler, M., Elhadad, M.: Enhancing unlexicalized parsing performance using a wide coverage lexicon, fuzzy tag-set mapping, and EM-HMM-based lexical probabilities. In: Proceedings of the 12th Conference of the European Chapter of the ACL, pp. 327–335 (2009)
4. Chen, J., Bangalorey, S., Rambow, O., Walkery, M.A.: Towards automatic generation of natural language generation systems. In: Proceedings of the Nineteenth International Conference on Computational Linguistics (2002)
5. Walker, M.A., Rambow, O., Rogati, M.: SPoT: a trainable sentence planner. In: Proceedings of the North American Meeting of the Association for Computational Linguistics (2001)
6. Paris, C., Colineau, N., Wilkinson, R.: Evaluations of NLG systems: common corpus and tasks or common dimensions and metrics. In: Proceedings of the Fourth International Natural Language Generation Conference, pp. 127–129 (2006)
7. de Araujo, R.P.A., de Oliveira, R.L., de Novais, E.M., Tadeu, T.D., Pereira, D.B., Paraboni, I.: SINotas: the evaluation of a NLG application. In: Proceedings of the International Conference on Language Resources and Evaluation, LREC (2010)
8. Chen, J., Bangalore, S., Rambow, O., Walker, M.: Towards automatic generation of natural language generation systems. In: International Conference on Computational Linguistics (2002)