Advances in Intelligent Systems and Computing 214

Fuchun Sun
Tianrui Li
Hongbo Li *Editors*

Knowledge Engineering and Management

Proceedings of the Seventh International Conference on Intelligent Systems and Knowledge Engineering, Beijing, China, Dec 2012 (ISKE 2012)

Springer

Advances in Intelligent Systems and Computing

Volume 214

Series Editor

J. Kacprzyk, Warsaw, Poland

For further volumes:
http://www.springer.com/series/11156

Fuchun Sun · Tianrui Li
Hongbo Li
Editors

Knowledge Engineering and Management

Proceedings of the Seventh International Conference on Intelligent Systems and Knowledge Engineering, Beijing, China, Dec 2012 (ISKE 2012)

Springer

Editors
Fuchun Sun
Hongbo Li
Department of Computer Science
and Technology
Tsinghua University
Beijing
People's Republic of China

Tianrui Li
School of Information Science
Southwest Jiaotong University
Chengdu
People's Republic of China

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-642-37831-7 ISBN 978-3-642-37832-4 (eBook)
DOI 10.1007/978-3-642-37832-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013942994

Printed on acid-free paper

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

Preface

This book is part of the Proceedings of the 7th International Conference on Intelligent Systems and Knowledge Engineering (ISKE2012) and the 1st International Conference on Cognitive Systems and Information Processing (CSIP2012) held in Beijing, China, during December 15–17, 2012. ISKE is a prestigious annual conference on Intelligent Systems and Knowledge Engineering with the past events held in Shanghai (2006, 2011), Chengdu (2007), Xiamen (2008), Hasselt, Belgium (2009), and Hangzhou (2010). Over the past few years, ISKE has matured into a well-established series of international conferences on Intelligent Systems and Knowledge Engineering and related fields over the world. CSIP 2012 is the first conference sponsored by Tsinghua University and Science in China Press, and technically sponsored by IEEE Computational Intelligence Society, Chinese Association for Artificial Intelligence. The aim of this conference is to bring together experts from different expertise areas to discuss the state of the art in cognitive systems and advanced information processing, and to present new research results and perspectives on future development. Both ISKE 2012 and CSIP 2012 provide academic forums for the participants to disseminate their new research findings and discuss emerging areas of research. It also creates a stimulating environment for the participants to interact and exchange information on future challenges and opportunities of intelligent and cognitive science research and applications.

ISKE 2012 and CSIP received 406 submissions in total from about 1020 authors in 20 countries (United States of American, Singapore, Russian Federation, Saudi Arabia, Spain, Sudan, Sweden, Tunisia, United Kingdom, Portugal, Norway, Korea, Japan, Germany, Finland, France, China, Argentina, Australia, and Belgium). Based on rigorous reviews by the Program Committee members and reviewers, among 220 papers contributed to ISKE 2012, high-quality papers were selected for publication in the proceedings with the acceptance rate of 58.4 %. The papers were organized in 25 cohesive sections covering all major topics of intelligent and cognitive science and applications. In addition to the contributed papers, the technical program includes four plenary speeches by Jennie Si (Arizona State University, USA), Wei Li (California State University, USA), Chin-Teng Lin (National Chiao Tung University, Taiwan, China), and Guoqing Chen (Tsinghua University, China).

As organizers of both conferences, we are grateful to Tsinghua University, Science in China Press, Chinese Academy of Sciences for their sponsorship, grateful to IEEE Computational Intelligence Society, Chinese Association for Artificial Intelligence, State Key Laboratory on Complex Electronic System Simulation, Science and Technology on Integrated Information System Laboratory, Southwest Jiaotong University, University of Technology, Sydney for their technical co-sponsorship.

We would also like to thank the members of the Advisory Committee for their guidance, the members of the International Program Committee and additional reviewers for reviewing the papers, and members of the Publications Committee for checking the accepted papers in a short period of time. Particularly, we are grateful to the publisher, Springer, for publishing the proceedings in the prestigious series of Advances in Intelligent Systems and Computing. Meanwhile, we wish to express our heartfelt appreciation to the plenary speakers, special session organizers, session chairs, and student helpers. In addition, there are still many colleagues, associates, and friends who helped us in immeasurable ways. We are also grateful to them all. Last but not the least, we are thankful to all authors and participants for their great contributions that made ISKE 2012 and CSIP 2012 successful.

December 2012

Fuchun Sun
Tianrui Li
Hongbo Li

Organizing Committee

General Chair:	Jie Lu (Australia)
	Fuchun Sun (China)
General Co-Chairs:	Yang Xu (China)
Honorary Chairs:	L. A. Zadeh (USA)
	Bo Zhang (China)
Steering Committee Chairs:	Etienne Kerre (Belgium)
	Zengqi Sun (China)
Organizing Chairs:	Xiaohui Hu (China)
	Huaping Liu (China)
Program Chairs:	Tianrui Li (China)
	Javier Montero (Spain)
Program Co-Chairs:	Changwen Zheng (China)
	Luis Martínez López (Spain)
Sessions Chairs:	Fei Song (China)
	Victoria Lopez (Spain)
Publications Chairs:	Yuanqing Xia (China)
	Hongming Cai (China)
Publicity Chairs:	Jiacun Wang (USA)
	Zheying Zhang (Finland)
	Michael Sheng (Australia)
	Dacheng Tao (Australia)
Poster Chairs:	Guangquan Zhang (Australia)
	Hongbo Li (China)
Program Chairs:	Tianrui Li (China)
	Javier Montero (Spain)

Members

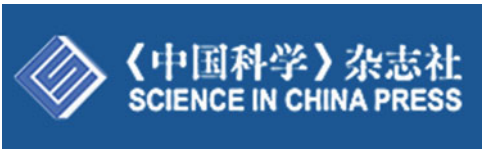
Abdullah Al-Zoubi (Jordan)
Andrzej Skowron (Poland)
Athena Tocatlidou (Greece)
B. Bouchon-Meunier (France)
Michael Sheng (Australia)
Mihir K. Chakraborty (India)
Mike Nachtegael (Belgium)
Mikhail Moshkov (Russia)

Benedetto Matarazzo (Italy)
Bo Yuan (USA)
Bo Zhang (China)
Cengiz Kahraman (Turkey)
Changwen Zheng
Chien-Chung Chan (USA)
Cornelis Chris (Belgium)
Dacheng Tao (Australia)
Davide Ciucci (Italy)
Davide Roverso (Norway)
Du Zhang (USA)
Enrico Zio (Italy)
Enrique Herrera-Viedma (Spain)
Erik Laes (Belgium)
Etienne E. Kerre (Belgium)
Francisco Chiclana (UK)
Francisco Herrera (Spain)
Fuchun Sun (China)
Gabriella Pasi (Italy)
Georg Peters (Germany)
Germano Resconi (Italy)
Guangquan Zhang (Australia)
Guangtao Xue (China)
Gulcin Buyukozkan (Turkey)
Guolong Chen (China)
Guoyin Wang (China)
H.-J. Zimmermann (Germany)
Huaping Liu (China)
Hongbo Li (China)
Hongjun Wang (China)
Hongming Cai (China)
Hongtao Lu (China)
I. Burhan Turksen (Canada)
Irina Perfilieva (Czech Republic)
Jan Komorowski (Sweden)
Janusz Kacprzyk (Poland)
Javier Montero (Spain)
Jer-Guang Hsieh (Taiwan, China)
Jesús Vega (Spain)
Jiacun Wang (USA)
Jianbo Yang (UK)
Jie Lu (Australia)
Jingcheng Wang (China)
Min Liu (China)
Peijun Guo (Japan)
Pierre Kunsch (Belgium)
Qi Wang (China)
Qingsheng Ren (China)
Rafael Bello (Cuba)
Richard Jensen (UK)
Ronald R. Yager (USA)
Ronei Marcos de Moraes (Brasil)
Ryszard Janicki (Canada)
S. K. Michael Wong (Canada)
Shaojie Qiao (China)
Shaozi Li (China)
Sheela Ramanna (Canada)
Su-Cheng Haw (Malaysia)
Suman Rao (India)
Sushmita Mitra (India)
Takehisa Onisawa (Japan)
Tetsuya Murai (Japan)
Tianrui Li (China)
Tzung-Pei Hong (Taiwan, China)
Ufuk Cebeci (Turkey)
Victoria Lopez (Spain)
Vilem Novak (Czech Republic)
Weiming Shen (Canada)
Weixing Zhu (China)
Wensheng Zhang (China)
Witold Pedrycz (Canada)
Wujun Li (China)
Xiaohui Hu (China)
Xianyi Zeng (France)
Xiaogang Jin (China)
Xiaoqiang Lu (China)
Xiaoyan Zhu (China)
Xiao-Zhi Gao (Finland)
Xuelong Li (China)
Xun Gong (China)
Yan Yang (China)
Yangguang Liu (China)
Yanmin Zhu (China)
Yaochu Jin (Germany)
Yasuo Kudo (Japan)
Yi Tang (China)

Jitender S. Deogun (USA)
Jouni Jarvinen (Finland)
Juan-Carlos Cubero (Spain)
Jun Liu (UK)
Jyrki Nummenmaa (Finland)
Koen Vanhoof (Belgium)
Krassimir Markov (Bulgaria)
Liliane Santos Machado (Brasil)
Lisheng Hu (China)
Luis Magdalena (Spain)
Luis Martinez López (Spain)
Lusine Mkrtchyan (Italy)
Madan M. Gupta (Canada)
Martine De Cock (Belgium)
Masoud Nikravesh (USA)
Yinglin Wang (China)
Yiyu Yao (Canada)
Yongjun Shen (Belgium)
Yuancheng Huang (China)
Yuanqing Xia (China)
Zbigniew Suraj (Poland)
Zbigniew W. Ras (USA)
Zengqi Sun (China)
Zheying Zhang (Finland)
Zhong Li (Germany)
Zhongjun He (China)
Zhongzhi Shi (China)
Bo Peng (China)
Fei Teng (China)

Sponsors Logo

Tsinghua University



Science in China Press

Chinese Academy of Sciences

Institute of Electrical and Electronics Engineers

Contents

Function Set-Valued Information Systems

Hongmei Chen, Tianrui Li, Chuan Luo, Junbo Zhang and Xinjiang Li

Abstract Set-valued Information System (SIS) aims at dealing with attributes' multi-values. It is an extension of the single-valued information system. In traditional SIS, the tolerance and dominance relations were introduced to deal with the relationship between objects. But the partial orders between the attributes values have not been taken into consideration. In this paper, a function set-valued equivalence relation is proposed first considering the applications of SIS. Then the properties of the function set-valued equivalence classes are analyzed. Furthermore, the definitions of approximations under the function SIS are presented and the properties w.r.t. approximations under the tolerance and dominance relations are investigated.

This work is supported by the National Science Foundation of China (Nos. 61175047, 61100117) and NSAF (No. U1230117), the Youth Social Science Foundation of the Chinese Education Commission (No. 11YJC630127), and the Fundamental Research Funds for the Central Universities (SWJTU11ZT08, SWJTU12CX091).

H. Chen (✉) · T. Li · C. Luo · J. Zhang · X. Li
School of Information Science and Technology, Southwest Jiaotong University, 610031 Chengdu, China
e-mail: hmchen@swjtu.edu.cn

T. Li
e-mail: trli@swjtu.edu.cn

C. Luo
e-mail: luochuan@my.swjtu.edu.cn

J. Zhang
e-mail: JunboZhang86@163.com

F. Sun et al. (eds.), *Knowledge Engineering and Management*,
Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_1,

1 Introduction

Granular Computing (GrC) proposed by Zadeh [1, 2] in 1989 has been successfully used in many areas, e.g., image processing, pattern recognition, and data mining. A granule is a chunk of knowledge made of different objects "drawn together by indistinguishability, similarity, proximity or functionality" [2]. Different levels of concepts or rules are induced in GrC. Rough set theory proposed by Pawlak in 1982 is considered as an important branch of GrC, which has been used to process inconsistent information [3].

In Traditional Rough Set theory (TRS), an attribute of one object in the information system can only have one value. But in real-life applications, attributes may have multi-values. For example, the degree of a person may be a combination of bachelor, master, and doctor. Set-valued systems are used to handle the case when objects' attribute have multi-values. In addition, there exists a lot of data in the information system which have uncertainty and incompleteness. Set-valued Information System (SIS) can also be used to process incomplete information. In this case, the missing value can be only one of the values in the value domain [4]. In [4], Guan defined the tolerance relation and the maximum tolerance block as well as the relative reduct based maximum tolerance block in SIS. Qian and Liang defined disjunctive and conjunctive SIS in [5] in view of different meaning of attributes' values and they studied the approach of attribute reduction and rule extraction in these two different SISs. Song and Zhang et al. defined the partly accordant reduction and the assignment reduction in an inconsistent set-valued decision information system [6, 7]. Huang et al. introduced a degree dominance relation to dominance set-valued intuitionist fuzzy decision tables and a degree dominance set-valued rough set model [8]. In [9, 10], Chen et al. introduced the probability rough sets into SIS and they studied the method for updating approximations incrementally under the variable precision set-valued ordered information while attributes' values coarsening and refining.

The equivalence relation is employed in TRS to deal with the relationship among objects and then equivalence classes induced by different equivalence relations form the elementary knowledge of the universe. Any set in the universe is described approximately by equivalence classes. Hence, a pair of certain sets, i.e., upper and lower approximations are used to describe the set approximately. The approximations partition the universe into three different regions, namely, positive, negative, and boundary regions. Then we can induce certain rules from the positive and negative regions and uncertain rules from the boundary region, i.e., a tree-way decision [11]. The SIS is a general model of single-valued information system. The equivalence relation is substituted by the tolerance relation and the dominance relation. Because the value of the missing data may be any value in the values' domain, the SIS has been used to process data in the incomplete information system. The tolerance and dominance relations in SIS consider multi-values of attributes. But the partial order among the attributes' values is not taken into consideration. In this paper, by analyzing the meaning of the tolerance and

dominance relations, a new relationship in SIS is proposed, i.e., a function set-valued equivalence relation. Then the properties of the function set-valued equivalence relation are discussed. Furthermore, a more general relationship, a general function relation, is investigated and its properties are analyzed.

The paper is organized as follows. In Sect. 2, basic concepts in SIS are reviewed. In Sect. 3, a function SIS is proposed. A function equivalence relation and its properties are discussed. Then the general function relation is investigated. The paper ends with conclusions and further directions in Sect. 4.

2 Preliminary

In this section, Some basic concepts in SIS are briefly reviewed [4–6].

Definition 21 Let $S = (U, A, V, f)$ be a SIS, where $U = \{x_1, x_2, \ldots, x_n\}$ is a non-empty finite set of objects, called the universe. $A = \{a_1, a_2, \ldots, a_l\}$ is a non-empty finite set of attributes. The element in A is called an attribute. $A = C \cup D$, $C \cap D = \emptyset$, C is the set of condition attributes and D is the set of decision attributes. $V = \{V_{a_1}, V_{a_2}, \ldots, V_{a_l}, a_i \in A\}$, $V_{a_i}(i = 1, 2, \ldots, l)$ is the domain of attribute $a_i(a_i \in A)$. V_C is the domain of condition attributes. V_D is the domain of decision attributes. $V = V_C \cup V_D$ is the domain of all attributes. $f : U \times C \to 2^{V_C}$ is a set-valued mapping.$f : U \times D \to V_D$ is a single-valued mapping. $f(x_i, a_l)$ is the value of x_i on attribute a_l. An example of SIS is shown in Table 1.

The tolerance and dominance relation have been proposed to deal with the different applications in SIS. For example, if two persons have the tolerance relation on the attribute "language", it means they can communicate through the common language. The tolerance relation is defined as follows.

Definition 22 Let $S = (U, A, V, f)$ be an SIS, for $B \subseteq C$, the tolerance relation in the SIS is defined as follows:

$$R_B^{\cap} = \{(y, x) \in U \times U | f(y, a) \cap f(x, a) \neq \emptyset (\forall a \in B)\} \tag{1}$$

Table 1 A set-valued information system

U	Price	Mileage	Size	Max-speed
1	{High, Medium}	High	Full	{High, Medium}
2	Low	{High, Medium}	Full	Medium
3	Medium	Low	Full	Medium
4	{Low, Medium}	{Low, Medium}	Compact	Low
5	High	{High, Medium}	Compact	{Low, Medium}
6	Medium	{Low, Medium}	Full	High
7	Low	High	Compact	High
8	High	Low	Compact	{Low, Medium}

The tolerance relation is reflexive and symmetric, but not transitive. Let $[x]_B^{\cap}$ denote the tolerance class of object x, where $[x]_B^{\cap} = \{y \in U \big| (y,x) \in R_B^{\cap}\} = \{y \in U | f(x,a_i) \cap f(y,a_i) \neq \emptyset, a_i \in B\}$.

Definition 23 Let $S = (U,A,V,f)$ be an SIS, $\forall X \subseteq U$, the approximations of X under the tolerance relation $R_C^{\cap}$ are defined as follows, respectively:

$$\underline{apr_B^T}(X) = \{x \in U \big| [x_i]_C^{\cap} \subseteq X\} \tag{2}$$

$$\overline{apr_B^T}(X) = \{x \in U \big| [x_i]_C^{\cap} \cap X \neq \emptyset\} \tag{3}$$

Definition 24 Given an SIS (U,A,V,f), for $B \supseteq C$, a dominance relation in the SIS is defined as follows:

$$\begin{aligned} R_B^{\subseteq} &= \{(y,x) \in U \times U | f(y,a) \supseteq f(x,a)(\forall a \in B)\} \\ &= \{(y,x) \in U \times U | y \succeq x\} \end{aligned} \tag{4}$$

$R_B^{\subseteq}$ is reflexive, dissymmetric, and transitive.

We denote $[x]_B^{\subseteq} = \{y \in U \big| (y,x) \in R_B^{\geq}\}$, $[x]_B^{\subseteq} = \{y \in U \big| (x,y) \in R_B^{\subseteq}\}$. $[x]_B^{\subseteq}$ $(x \in U)$ are granules of knowledge induced by the dominance relation, which are the set of objects dominating x. The dominance relation in SIS means in some degree, the more values in an attribute of one object the more the probability that it dominates the other objects. It considers the cardinality of the attributes' values.

Definition 25 Let $S = (U,A,V,f)$ be an SIS, $\forall X \subseteq U$, the approximations of X under the dominance relation $R_C^{\supseteq}$ are defined as follows, respectively:

$$\underline{apr_B^D}(X) = \{x \in U \big| [x_i]_C^{\supseteq} \subseteq X\} \tag{5}$$

$$\overline{apr_B^D}(X) = \{x \in U \big| [x_i]_C^{\supseteq} \cap X \neq \emptyset\} \tag{6}$$

3 Function Set-Valued Information Systems

Multi-values of attributes in SIS give more information than single-valued information system. In real-life applications, a partial order may exist in the attribute values, i.e., for $V_{a_i} = \{v_{a_1}^1, v_{a_1}^2, \cdots, v_{a_1}^k\}$ $(k = |V_{a_i}|)$, $v_{a_1}^1 \preceq v_{a_1}^2 \preceq \cdots \preceq v_{a_1}^k$ is the partial order among the attribute values on the attribute a_i. In Definitions 22 and 24, the multi-value property of attributes is taken into consideration but the partial order in the attribute values is neglected. Then, we present a new definition, namely, a function set-valued equivalence relation.

Definition 31 Let $S = (U, A, V, f)$ be a SIS, $V_{a_i} = \{v_{a_1}^1, v_{a_1}^2, \ldots, v_{a_1}^k\}(k = |V_{a_i}|)$ is the attribute value domain of attribute a_i, where $v_{a_1}^i (1 \leq i \leq k)$ *s.t.* the $v_{a_1}^1 \preceq v_{a_1}^2 \preceq \ldots \preceq v_{a_1}^k$. Then an attribute function F_{a_i} is defined as $F_{a_i} : V_{a_i} \rightarrow P(V_{a_i})$, which is a mapping from V_{a_i} to the power set of V_{a_i}.

The definition of F_{a_i} is given by the requirement of real-life applications. Usually, F_{a_i} may be maximum, minimum, average of the set-valued. For example, the degree a person awarded may be bachelor, master, doctor; the employer in universities may be more interesting in the person with the highest degree. Let $F(f(x, a_i))$ denote the value of the function F_{a_i} on the attribute value $f(x_i, a_l)$. Then the function equivalence relation is defined as follows:

Definition 32 Let $S = (U, A, V, f)$ be an SIS, the function equivalence relation is defined as follows:

$$R_B^F = \{(y, x) \in U \times U | F(f(x, a_i)) = F(f(y, a_i)), a_i \in B\} \tag{7}$$

Let $[x]_B^F$ denote the function equivalence class of object x, where $[x]_B^F = \{y \in U \big| (y, x) \in R_B^F\}$.

Property 31 For R_B^F, the following hold:

1. Reflexive: $\forall x \in U$, $xR_B^F x$;
2. Symmetric: $\forall x, y \in U$, if $xR_B^F y$, then $yR_B^F x$;
3. Transitive: $\forall x, y, z \in U$, if $xR_B^F y$ and $yR_B^F z$, then $xR_B^F z$;

$\bigcup_{i=1}^{|U|} [x_i]_B^F = U$ and $[x_i]_B^F \cap [x_j]_B^F = \emptyset$, i.e., U/R_B^F forms a partition of the university.

For any $B \subseteq C$, $[x_i]_C^F \subseteq [x_i]_B^F$. The tolerance, dominance, and function equivalence relation meet the different requirements of real-life applications. In the following, we compare these three different relations defined in SIS. Let $[x_i]_B^*$ denote the classes induced by different relations in a SIS and $\widetilde{I^*} = \{[x_i]_B^* \big| x_i \in U\}$ denote a classes cluster set, where the superscript * may be replaced by $\cap$, $\supseteq$, F and $\widetilde{I^\cap}$, $\widetilde{I^\supseteq}$, $\widetilde{I^F}$ denote tolerance, dominance, and function equivalence classes, respectively. Liang et al. defined the granularity of knowledge under the tolerance relation in incomplete information system [12]. Similarly, the granularity and the rough entropy of B in an SIS is defined as follows:

Definition 33 Let $S = (U, A, V, f)$ be a SIS, $[x_i]_B^* (x_i \in U, i = 1, 2, \ldots, |U|)$ is the granule induced by different relation in a set-valued information system, $\forall B \subseteq C \subseteq A$, where the superscript $*$ may be substituted by $\cap$, $\supseteq$, F and the granularity of B is defined as follows:

$$GK^*(B) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{\left|[x_i]_B^*\right|}{|U|} \tag{8}$$

Definition 34 Let $S = (U, A, V, f)$ be an SIS, $[x_i]_B^*(x_i \in U, i = 1, 2, \ldots, |U|)$ is granule induced by different relation in a set-valued information system, $\forall P \subseteq C \subseteq A$. Then the rough entropy of P is defined as follows:

$$E_r^*(P) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{1}{\left|[x_i]_B^*\right|} \tag{9}$$

For convenience, we define a partial order binary relation $\underline{\prec}$ first.

$I^* \underline{\prec} I^{\Delta} \Leftrightarrow$ *if* $\forall x_i \in U, \exists [x_i]_B^* \subseteq [x_i]_B^{\Delta}$, where $[x_i]_B^* \in \widetilde{I^*}$ and $[x_i]_B^{\Delta} \in \widetilde{I^{\Delta}}$.

Note: The superscript Δ may also be replaced by $\cap$, $\supseteq$, F. Then the following properties hold:

Property 32 For $[x_i]_B^*$ and $\widetilde{I^*}$, the following properties hold:

1. For any $B_1 \subset B_2 \subseteq C$, $[x_i]_{B_2}^F \subseteq [x_i]_{B_1}^F \subseteq [x_i]_C^F$;
2. For any $B \subseteq C$, $[x_i]_B^F \subseteq [x_i]_B^{\cap}$;
3. For any $B \subseteq C$, $I^F \underline{\prec} \widetilde{I^{\cap}}$;
4. For any $B \subseteq C$, $GK^{\supseteq}(B) \leq GK^{\cap}(B)$;
5. For any $B \subseteq C$, $E_r^{\cap}(P) \leq E_r^{\supseteq}(P)$.

Proof Proof 1. $\forall x_j \in [x_i]_{B_2}^F$, $x_j R_{B_2}^F x_i$, i.e., $F(f(x_i, a_i)) = F(f(x_j, a_i)), \forall a_i \in B_2$. Since $B_1 \subset B_2$, then $x_j \in [x_i]_{B_1}^F$. On the other hand, if $\forall x_j \in [x_i]_{B_1}^F$, $x_j R_{B_1}^F x_i$, i.e., $F(f(x_i, a_i)) = F(f(x_j, a_i))$, $\forall a_i \in B_1$. Because $B_1 \subset B_2$, then if $\exists F(f(x_i, a_i)) \neq F(f(x_j, a_i))$, $\forall a_i \in B_2 - B_1$. Then $x_j \not R_{B_2}^F x_i$, i.e., $x_j \notin [x_i]_{B_1}^F$. Therefore, we have $[x_i]_{B_2}^F \subseteq [x_i]_{B_1}^F$. Consequently, for any $B_1 \subset B_2 \subseteq C$, $[x_i]_C^F \subseteq [x_i]_{B_2}^F \subseteq [x_i]_{B_1}^F$. 2. $\forall x_j \in [x_i]_C^F$, $F(f(x_i, a_i)) = F(f(x_j, a_i)), a_i \in B$. Because $F_{a_i} : V_{a_i} \to P(V_{a_i})$, then $f(x_i, a_i) \cap f(x_j, a_i) \neq \emptyset$, i.e., $\forall x_j \in [x_i]_C^{\cap}$. $\forall x_j \in [x_i]_C^{\cap}$, we have $f(x_i, a_i) \cap f(x_j, a_i) \neq \emptyset$, if $\exists F(f(x_i, a_i)) \neq F(f(x_j, a_i)), a_i \in B$, $x_j \notin [x_i]_C^F$. Therefore, $[x_i]_B^F \subseteq [x_i]_B^{\cap}$. Consequently, for any $B \subseteq C$, $[x_i]_B^F \subseteq [x_i]_B^{\cap}$. Since 2 is always true, then 3, 4, and 5 hold.

The tolerance, dominance, and function equivalence relations are used to deal with different cases in a SIS. In most cases, three kinds of relations may coexist on different attributes. In the following, a general relation is given.

Definition 35 Let $S = (U, A, V, f)$ be a SIS, where $B_T \subseteq C$, $B_D \subseteq C$, $B_F \subseteq C$, $B_T \cap B_D \cap B_F = \emptyset$, $B_T \cup B_D \cup B_F = B \subseteq C$, the general function relation is defined as follows:

$$\begin{aligned} R_B^G = \{(y,x) \in U \times U | f(x,a_t) \cap f(y,a_t) \neq \emptyset \wedge f(x,a_d) \subseteq f(y,a_d) \\ \wedge F(f(x,a_f)) = F(f(y,a_f)), a_t \in B_T, a_d \in B_D, a_f \in B_F\} \end{aligned} \tag{10}$$

Let $[x]_C^G$ denote the function equivalence class of object x, where $[x]_B^G = \{y \in U \big| (y,x) \in R_B^G\}$.

Property 33 For R_B^G,

1. Reflexive: $\forall x \in U$, $xR_B^G x$;
2. Dissymmetric: $\forall x, y \in U$, if $xR_B^F y$, then $y\overline{R_B^G}x$;
3. Non-transitive.

$\bigcup_{i=1}^{|U|} [x_i]_B^G = U$and $[x_i]_B^G \cap [x_j]_B^G = \emptyset$, i.e., U/R_B^G forms a covering of the university.

Then, the approximations based on the general function relation are defined as follows:

Definition 36 Let $S = (U, A, V, f)$ be a SIS, $\forall X \subseteq U$, the lower and the upper approximations of X under the general function relation R_C^G are defined as follows, respectively:

$$\underline{apr_B^G}(X) = \{x \in U \big| [x_i]_C^G \subseteq X\} \tag{11}$$

$$\overline{apr_B^G}(X) = \{x \in U \big| [x_i]_C^G \cap X \neq \emptyset\} \tag{12}$$

Based on approximations of X, one can partition the universe U into three disjoint regions, i.e., the *positive region* $POS_B^G(X)$, the *boundary region* $BNR_B^G(X)$, and the *negative region* $NEG_B^G(X)$:

$$POS^G(X) = \underline{apr_B^G}(X); \tag{13}$$

$$BNR^G(X) = \overline{apr_B^G}(X) - \underline{apr_B^G}(X); \tag{14}$$

$$NEG^G(X) = U - \overline{apr_B^G}(X). \tag{15}$$

We can extract certain rules from the position and negative regions and possible rules from the boundary region too.

Property 34 For $\underline{apr_B^G}(X)$ and $\overline{apr_B^G}(X)$, the following hold:

$$\underline{apr_B^T}(X) \subseteq \underline{apr_B^G}(X);$$

$$\overline{apr_B^D}(X) \subseteq \overline{apr_B^G}(X) \subseteq \overline{apr_B^T}(X);$$

$$If\ A \subseteq B \subseteq C, then\ \underline{apr_A^G}(X) \subseteq \underline{apr_B^G}(X);$$

$$If\ A \subseteq B \subseteq C, then\ \overline{apr_B^G}(X \subseteq \overline{apr_A^G}(X);$$

Proof (1) From Property 32, for any $B \subseteq C$, $[x_i]_B^F \subseteq [x_i]_B^{\cap}$, then $[x_i]_B^G \subseteq [x_i]_B^{\cap}$. Therefore, we have $\underline{apr_B^T}(X) \subseteq \underline{apr_B^G}(X)$.

4 Conclusions

In this paper, the partial order among the multi-values in an SIS is taken into consideration. The function equivalence relation is proposed. Then the properties of function equivalence classes are analyzed. Furthermore, a general function relation is defined in the SIS. Then, approximations and its properties under the general function relation are discussed. In the future work, we will study the attribute reduction under the general function relation in the SIS.

References

1. Zadeh LA (1997) Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 19(1):111–127
2. Zadeh LA (2008) Is there a need for fuzzy logic? Inf Sci 178:2751–2779
3. Pawlak Z, Skowron A (2007) Rudiments of rough sets. Inf Sci 177:3–27
4. Guan YY, Wang HK (2006) Set-valued information systems. Inf Sci 176(17):2507–2525
5. Qian YH, Dang CY, Liang JY, Tang DW (2009) Set-valued ordered information systems. Inf Sci 179(16):2809–2832
6. Song XX, Zhang WX (2009) Knowledge reduction in inconsistent set-valued decision information system. Comput Eng Appl 45(1):33–35 (In chinese)
7. Song XX, Zhang WX (2006) Knowledge reduction in set-valued decision information system. In: Proceedings of 5th international conference on rough sets and current trends in computing (RSCTC 2006), 4259 LNAI, pp 348–357
8. Huang B, Li HX, Wei DK (2011) Degree dominance set-valued relation-based RSM in intuitionistic fuzzy decision tables. In: Proceedings of the international conference on internet computing and information services, pp 37–40
9. Chen HM, Li TR, Zhang JB (2010) A method for incremental updating approximations based on variable precision set-valued ordered information systems. In: Proceedings of the 2010 IEEE international conference on granular computing (GrC2010), pp 96–101

10. Zou WL, Li TR, Chen HM, Ji XL (2009) Approaches for incrementally updating approximations based on set-valued information systems whileattribute values' coarsening and refining. In: Proceedings of the 2009 IEEE international conference on granular computing (GrC2009), pp 824–829
11. Yao YY (2010) Three-way decisions with probabilistic rough sets. Inf Sci 180(7):341–353
12. Liang JY, Shi ZZ (2004) The information entropy, rough entropy and knowledge granulation in rough set theory. Int J Uncertainty Fuzziness Knowl Based Syst 12(1):37–46

Neutrality in Bipolar Structures

Javier Montero, J. Tinguaro Rodriguez, Camilo Franco, Humberto Bustince, Edurne Barrenechea and Daniel Gómez

Abstract In this paper, we want to stress that bipolar knowledge representation naturally allows a family of middle states which define as a consequence different kinds of bipolar structures. These bipolar structures are deeply related to the three types of bipolarity introduced by Dubois and Prade, but our approach offers a systematic explanation of how such bipolar structures appear and can be identified.

Keywords Bipolarity · Compatibility · Conflict · Ignorance · Imprecision · Indeterminacy · Neutrality · Symmetry

J. Montero (✉) · J. T. Rodriguez
Faculty of Mathematics, Complutense University, Madrid, Spain
e-mail: monty@mat.ucm.es

J. T. Rodriguez
e-mail: jtrodrig@mat.ucm.es

C. Franco
Centre for Research and Technology of Agro-Environmental and Biological Sciences, Tras-os-Montes e Alto Douro University, Vila Real, Portugal
e-mail: camilo@utad.pt

H. Bustince · E. Barrenechea
Departamento de Automática y Computación, Universidad Pública de Navarra, Pamplona, Spain
e-mail: bustince@unavarra.es

E. Barrenechea
e-mail: edurne.barrenechea@unavarra.es

D. Gómez
School of Statistics, Complutense University, Madrid, Spain
e-mail: dagomez@estad.ucm.es

F. Sun et al. (eds.), *Knowledge Engineering and Management*,
Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_2,

1 Introduction

Dubois and Prade [8–10] distinguished between three types of bipolarity, named Type I, Type II, and Type III bipolarity, but in our opinion a unified approach for such a classification is missing.

Our approach in this paper focuses on how such bipolar models are being built in knowledge representation. Our main point is that bipolarity appears whenever two opposite arguments are taken into account, being the key issue how intermediate *neutral* stages are being generated from such an opposition. Such intermediate stages can be introduced from different circumstances and generate a different relationship with the basic two opposite arguments.

As pointed out in [17], different roles should be associated to different structures, and different structures justify different concepts. In particular, we claim that a semantic approach to bipolarity will allow different *neutral* stages between the two extremes under consideration, depending on the nature of these two extreme values. In this way, concepts such as *imprecision*, *indeterminacy*, *compatibility,* and *conflict* between poles will be distinguished. These *neutral* stages are somehow frequently understood as different forms of a heterogeneous *ignorance*. For example, ignorance—lack of information—is sometimes confused with symmetry in decision making—difficulties to choose between two poles despite clear available information. The lack of information can be based on simple imprecision or a deeper conceptual problem, whenever the two considered poles are not enough to fully explain reality. Similarly, a decision maker cannot be able to choose a unique pole when both poles simultaneously hold or when a conflict is being detected (random decision can be acceptable in the first case but not in the second place, where each pole can be rejected because of different arguments).

We should remind that Dubois and Prade [8] classify bipolarity in terms of the nature of the scales that are used and the relation between positive and negative information, differentiating as a consequence two types of bipolar scales: *univariate bipolar* and *bivariate unipolar* scales. Univariate bipolarity was associated in [8] and [14] to a linearly ordered set L in which the two ends are occupied by the poles $+$ and $-$, and certain *middle* value 0 might separate positive evaluations from the negative evaluations. An object is evaluated by means of a single value on L. On the other hand, bivariate unipolar scales admit in [8] positive and negative information to be measured separately by means of two unipolar scales, each one being occupied by a pole and a *neutral* state 0 that somehow appears in between both scales, but not as a middle value as in type I bipolarity. An object can receive two evaluations, which can be perceived as neither positive nor negative, as well as both positive and negative. This is the way type II and type III bipolarities are introduced in [8].

In the following sections, we shall offer an explanation to all those neutral stages between poles which will allow a systematic characterization of each different kind of bipolarity. In fact, we will see that the nature of the middle stage can be used to identify which bipolarity we are dealing with, although it should be

acknowledged that complex problems cannot be explained by means of a unique bipolar structure: different bipolarities will simultaneously appear in practice, showing that several semantics might coexist. Our approach will lead to the same differentiation proposed by Dubois and Prade, but giving a key role to the underlying structure-building process generated from the opposite poles and their associated semantic, that will produce in each circumstance a specific neutral middle stage.

2 Building Type I Bipolar Fuzzy Sets

Type I bipolarity assumes a scale that shows tension between a pole and its own negation. It suggests gradualness within a linear scale, and the middle intermediate value simply represents a scale *point of symmetry* (not that both poles hold). This linear structure is the characteristic property of type I bipolarity, and a single value in this scale simultaneously gives the distance to each pole. But such a symmetry value cannot be confused with compatibility or ignorance; it does not properly represent a new concept.

For example, meanwhile "tall" and "short" are viewed as two opposite grades of a unique "tallness" concept, such a "tallness" is being modeled as a type I bipolarity. In this case, we neither estimate "tallness" or "shortness", but "height".

Neutrality within two type I bipolar poles can more properly be associated to the unavoidable imprecision problem (uncertainty about the right value). *Imprecision* represents a very particular kind of *ignorance*, indeed an epistemic state different than both poles. When a decision maker has no information about the exact value, imprecision is maximum and the more information we get the more accurate we are.

Imprecision is the characteristic neutral state in type I bipolarity.

3 Building Type II Bipolar Fuzzy Sets

On the contrary, type II bipolarity requires the existence of two dual (perhaps antagonistic) concepts, related because they refer to a common concept but containing different information. Distance to one pole cannot be deduced from the distance to the other pole, so two separate evaluations are needed for each pole, although they share a common nature. Poles are not viewed as the extreme values of a unique gradualness scale like in type I bipolarity. Depending on the nature of such a duality, we may find different neutral intermediate states with differentiated meaning or semantics.

For example, two opposite concepts not necessarily cover the whole universe of discourse (as shown in [2] within a classification framework when compared to [19]). An object can neither fulfill a concept nor its opposite, a situation that by

definition cannot happen within type I bipolarity, where negation connects both poles. If classification poles are too distant, this situation suggests the need of some additional classification concept [1], and meanwhile we do not find such an additional concept, we have to declare an *indeterminacy* that is another kind of neutral *ignorance*, different in nature to previous *imprecision* (see [7]). This is the case when the dual of the concept is strictly contained in its negation (for example, "very short" is strictly contained in "not tall"). Poles do not cover the whole space of possibilities. There exists a region where none of both poles hold.

Alternatively, negation can be implied by its dual concept, in such a way that both poles overlap (as "not very tall" overlaps "not very short", see [6]). In this case, type II bipolarity generates a different specific intermediate concept that cannot be associated to *indeterminacy*, but to a different kind of *neutrality*. We are talking about a simultaneous verification of both poles which is not a point of symmetry. In type I bipolarity, symmetry refers to a situation *in between* poles; meanwhile, in this type II bipolarity, *compatibility* means that both poles simultaneously hold.

In this way, we find that type II bipolarity may generate two different neutral intermediate concepts, depending on the semantic relation between poles. While in type I bipolarity one of the poles is precisely the negation of the other, in type II bipolarity, two alternative type II bipolarities may appear (but they will not simultaneously appear).

4 Building Type III Bipolar Fuzzy Sets

Type III bipolarity refers to negative and positive pieces of information (see [11]), implying the existence of two families of arguments, that should be somehow aggregated. Poles here represent like bags of arguments. Poles are not direct arguments like in type I and type II bipolarity.

Type III bipolarity suggests a different construction than the one used for type I or type II bipolarity. In such type I or type II bipolarity, we start from the opposition between two extreme values within a single characteristic or between two dual poles, but in both cases assuming one single common concept. A concept generates its negation or some kind of dual concept, and their description may need one single value or two values, which can be directly estimated.

Type III bipolarity appears like a second degree bipolar type II structure, being both poles complex concepts like *positive–negative* or *good–bad*, for example, each pole needing a specific description or decomposition to be understood.

Type III bipolarity is essentially more complex than type II bipolarity, for example when we are asked to list on one side "positive" arguments and "negative" arguments on the other side.

This is the standard situation in multi-criteria decision making. It is in this context where *conflict* can naturally appear as another neutrality stage, besides *imprecision*, *compatibility,* and *indeterminacy*. In a complex problem, of course we can find at the same time strong arguments supporting both poles (see, e.g., [18]).

Such a *conflict* stage is natural within type III bipolarity, but of course it may happen that those arguments as a whole do not suggest a complete description of reality, suggesting *indeterminacy*, similarly to type II *indeterminacy*. Notice that in type II bipolarity *compatibility* can appear, but *conflict* should not be expected when dealing with a pair of extreme values coming from simple (1-dimensional) argument. Such a *conflictive* argument belongs to the type III bipolarity context.

It is worth noticing that meanwhile type I and type II bipolarity are built up from symmetrical poles, type III bipolarity is built from asymmetrical poles (see [8, 9]).

The key issue is the different semantic relation between poles. A different semantic relation produces a different intermediate *neutral* state.

Finally, it must be pointed out that in this third more complex problem all previous situations might be implied. Underlying criteria can define a conflict, but they can also define *indeterminacy* situation, resembling the type II *indeterminacy* as already pointed out (too poor descriptions suggesting *ignorance*), and in addition, each underlying criteria is subject to a symmetrical bipolarity framework, allowing *imprecision* or *compatibility*.(besides type II *indeterminacy*). For example, the semantic relation between "very good" and "very bad" is the same relation as "very tall" and "very short", allowing when something is neither "very good" or "very bad" a similar *indeterminacy* to the one that appears when something is neither "very tall" or "very short".

But *conflict* is essentially a type III bipolarity performance.

5 Building General Bipolarities

From the above comments it is clear that some bipolar problems require a quite complex structure.

Type I bipolarity use to imply the existence of an *imprecision* state besides a possible *point of symmetry*.

Type II bipolar implies two potential different intermediate states (*indeterminacy* and *compatibility*).

Type III bipolarity implies the possibility of a *conflict* stage, but previous intermediate stages can also appear associated to each underlying criteria, if not directly.

A general bipolarity model should allow all those four semantic neutral states between poles beside the non semantic *points of symmetry* that may appear within type I bipolarity. In principle, a general bipolar representation should be prepared to simultaneously deal with and evaluate all these states.

The semantic argument is anyway needed to distinguish between different bipolarities, and to produce structured type-2 fuzzy sets [16], as proposed in [17].

It is also interesting to realize that this semantic approach allows an alternative explanation to Atanassov's intuitionistic fuzzy sets (see [3–5]), realizing how different examples given by this author can be explained by means of different semantics and therefore different bipolarities.

6 Conclusion

When poles are defined within a binary context, in terms of a single concept and its *negation* (e.g., "tall" and "not tall"), a linear order of gradation states is allowed around the *point of symmetry*, simply meaning equal distance to both concepts within a linear order, like and in Probability Theory [15] or Fuzzy Sets [20] (see also [12]). An object is associated *to some extent* to both poles by means of a unique family of intermediate *gradation* states. This is the framework for type I bipolarity.

In type II bipolarity, we have dual concepts as poles. These two dual poles can overlap (like "more or less tall" and "more or less short") or they can create a region that is far away from poles (like "very tall" and "very short"). Anyway, opposite poles are not complementary, so two different situations can be generated. If a pole is much smaller than the negation of the opposite pole, *indeterminacy* will be natural. If a pole is much bigger than the negation of the opposite pole, *compatibility* will naturally appear.

Type III bipolarity is the natural framework for multicriteria decision making (see [13]). Poles are complex concepts and need a multicriteria description, and of course different criteria can produce a *conflict*. But also each one of the other situations can appear within each simple criterion. In addition, such a multicriteria description can be so poor that a strict *ignorance* can appear.

Acknowledgments This research has been partially supported by the Government of Spain, grants TIN2009-07901, TIN2012-32482 and TIN2010-15055.

References

1. Amo A, Gomez D, Montero J, Biging G (2001) Relevance and redundancy in fuzzy classification systems. Mathware and Soft Computing 8:203–216
2. Amo A, Montero J, Biging G, Cutello V (2004) Fuzzy classification systems. Eur J Oper Res 156:495–507
3. Atanassov KT (1983) Intuitionistic fuzzy sets. VII ITKR's Session, Sofia, deposed in Central Science-Technical Library of Bulgarian academy of Science, 1697/84 (in Bulgarian)
4. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20:87–96
5. Atanassov KT (2005) Answer to D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk and H. Prade's paper "Terminological difficulties in fuzzy set theory—the case of "Intuitionistic fuzzy sets". Fuzzy Sets Syst 156:496–499
6. Bustince H, Fernandez J, Mesiar R, Montero J, Orduna R (2010) Overlap functions. Nonlinear Analysis 72:1488–1499
7. Bustince H, Pagola M, Barrenechea E, Fernandez J, Melo-Pinto P, Couto P, Tizhoosh HR, Montero J (2010) Ignorance functions. An application to the calculation of the threshold in prostate ultrasound images, Fuzzy sets and Systems 161:20–36
8. Dubois D, Prade H (2008) An introduction to bipolar representations of information and preference. Int J Intell Syst 23:866–877
9. Dubois D, Prade H (2009) An overview of the asymmetric bipolar representation of positive and negative information in possibility theory. Fuzzy Sets Syst 160:1355–1366

10. Dubois D, Prade H Gradualness uncertainty and bipolarity: making sense of fuzzy sets. In: Fuzzy sets and Systems (in press)
11. Franco C, Montero J, Rodriguez JT (2011) On partial comparability and preference-aversion models. In: Proceedings ISKE 2011, Shangai, China, 15–17 Dicimbre 2011
12. Goguen J (1967) L-fuzzy sets. Journal of Mathematical Analysis and Applications 18:145–174
13. Gonzalez-Pachón J, Gómez D, Montero J, Yáñez J (2003) Soft dimension theory. Fuzzy Sets Syst 137:137–149
14. Grabisch M, Greco S, Pirlot M (2008) Bipolar and bivariate models in multi-criteria decision analysis: descriptive and constructive approaches. Int J Intell Syst 23:930–969
15. Kolmogorov AN (1956) Foundations of the theory of probability. Chelsea Publishing Co., New York
16. Mendel JM (2007) Advances in type-2 fuzzy sets and systems. Inf Sci 177:84–110
17. Montero J, Gomez D, Bustince H (2007) On the relevance of some families of fuzzy sets. Fuzzy Sets Syst 158:2429–2442
18. Roy B (1991) The outranking approach and the foundations of ELECTRE methods. Springer, Berlin
19. Ruspini EH (1969) A new approach to clustering. Inf Control 15:22–32
20. Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353

Multirate Multisensor Data Fusion Algorithm for State Estimation with Cross-Correlated Noises

Yulei Liu, Liping Yan, Bo Xiao, Yuanqing Xia and Mengyin Fu

Abstract This paper is concerned with the optimal state estimation problem under linear dynamic systems when the sampling rates of different sensors are different. The noises of different sensors are cross-correlated and coupled with the system noise of the previous step. By use of the projection theory and induction hypothesis repeatedly, a sequential fusion estimation algorithm is derived. The algorithm is proven to be optimal in the sense of Linear Minimum Mean Square Error(LMMSE). Finally, a numerical example is presented to illustrate the effectiveness of the proposed algorithm.

Keywords State estimation · Data fusion · Cross-correlated noises · Asynchronous multirate multisensor

1 Introduction

Estimation fusion, or data fusion for estimation, is the problem of how to best utilize useful information contained in multiple sets of data for the purpose of estimating a quantity, e.g., a parameter or process [1]. It originated in the military field, and is now widely used in military and civilian fields, e.g., target tracking and localization, guidance and navigation, surveillance and monitoring, etc., due to its improved estimation accuracy, enhanced reliability, and survivability, etc.

Y. Liu (✉) · L. Yan · B. Xiao · Y. Xia · M. Fu
Key Laboratory of Intelligent Control and Decision of Complex Systems,
School of Automation, Beijing Institute of Technology,
100081 Beijing, China
e-mail: lyulei0929@gmail.com

L. Yan
e-mail: liping.yan@gmail.com

F. Sun et al. (eds.), *Knowledge Engineering and Management*,
Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_3,

Most of the earlier works were based on the assumption of cross-independent sensor noises. Bar-Shalom [2], Chong et al. [3], Hashmipour et al. [4] proposed several optimal state estimation algorithms based on Kalman filtering, respectively, in which all sensor measurements are fused by use of centralized fusion structure. In the practical applications, most of multisensor systems often have the correlated noises when the dynamic process is observed in a common noisy environment [5]. Moreover, because most of the real systems are described in continuous forms, discretization is necessary when to get the state estimation on line, and in the process, the system noise and the measurement noises are shown to be coupled. Of course, the centralized filter can still be used for the systems with correlated noises as it is still optimal in the sense of LMMSE. However, the computation and power requirements are too huge to be practical.

Hence, a few pieces of work deal with coupled sensor noises. Duan proposed one systematic way to handle the distributed fusion problem based on a unified data model in which the measurement noises across sensors at the same time may be correlated [6]. Song also dealt with the state estimation problem with cross-correlated sensor noises, and proved that under a mild condition it is optimal [5]. A small amount of papers consider the coupled sensor noises and the correlation between sensor noises and system noises. Xiao et al. [7] considered the two kinds of correlations by augmentation and the computation is complex.

In all the papers mentioned above, the sampling rates of different sensors are the same. Based on the multi-sensor dynamic system in which different sensors observe the same target state with different sampling rates, Yan et al. [8] put forward a kind of optimal state estimation algorithm. The algorithm has stronger feasibility and practicality than the traditional state fusion algorithm, but it does not take the correlations of noises into account. Shi et al. [9] discussed the estimation when multisensors have multirate asynchronous sampling rates. However, it does not consider the sensor-correlations either.

In this paper, when the noises of different sensors are cross-correlated and when they are also coupled with the system noise of the previous step, also, when the sampling rates of different sensors are different, by use of the projection theory, a sequential algorithm is formulated. We analyzed the performance of the algorithm, and it is shown to be optimal in the sense of LMMSE.

The paper is organized as follows. In Sect. 2, the problem formulation is presented. Section 3 describes the optimal state estimation algorithm. Section 4 is the simulation results and Sect. 5 draws the conclusion.

2 Problem Formulation

Consider the following generic linear dynamic system,

$$x(k+1) = A(k)x(k) + w(k), k = 0, 1, \ldots \tag{1}$$

$$z_i(k_i) = C_i(k_i)x(k_i) + v_i(k_i), i = 1, 2, \ldots, N \tag{2}$$

where, $x(k) \in R^n$ is the system state, $A(k) \in R^{n\times n}$ is the state transition matrix, and $w(k)$ is the system noise and is assumed to be Gaussian distributed with zero mean and variance being $Q(k)$, where $Q(k) \geq 0$. $z_i(k_i) \in R^{m_i}$ is the measurement of sensor i at time k_i. Assume the sampling rate of sensor i is S_i, and $S_i = S_1/n_i$, where n_i is known positive integers. Without loss of generality, let the sampling period of sensor 1 be the unit time, that is, $k_1 = k$. Then, sensor i has measurement at sampling point $n_i k$, in other words, $k_i = n_i k$. $C_i(k_i) \in R^{m_i\times n}$ is the measurement matrix. Measurement noise $v_i(k_i)$ is zero-mean and is Gaussian distributed with variance being $R(k)$, and

$$E\{w(k_i - 1)v_i^T(k_i)\} = S_i(k_i) \tag{3}$$

From the above equation, we can see that the measurement noises are coupled with the previous step system noise. Namely, $v_i(k_i)$ is correlated with $w(k_i - 1)$ at time $k = 0, 1, \ldots, i = 1, 2, \ldots, N$. If different sensors have measurements at the same time, their measurement noises are cross-correlated, i.e., $v_i(k_i)$ and $v_j(l_j)$ are coupled when $k_i = l_j$. That is, $E\{v_i(k_i)v_j^T(k_i)\} = R_{ij}(k_i) \neq 0$ where $i, j = 1, 2, \ldots, N$. For simplicity, denote $R_i(k_i) \triangleq R_{ii}(k_i) > 0, i = 1, 2, \ldots, N$.

The initial state $x(0)$ is independent of $w(k)$ and $v_i(k_i)$, where $k = 1, 2, \ldots$, $i = 1, 2, \ldots, N$, and is assumed to be Gaussian distributed with

$$\begin{cases} E\{x(0)\} = x_0 \\ cov\{x(0)\} = E\{[x(0) - x_0][x(0) - x_0]^T\} = P_0 \end{cases} \tag{4}$$

where $cov\{x(0)\}$ means the covariance of $x(0)$.

It can be seen from the above description that sensor i will participate in the fusion process at time $n_i k$. Generally speaking, assume there are p sensors that have measurements at time k with measurements $z_{i_1}(k), z_{i_2}(k), \ldots, z_{i_p}(k)$. Then, to generate the optimal state estimate of x(k) in the measurement update step, the above p sensors shall be fused. The estimation of $x(k)$ is the information fusion of the above p sensors.

3 Optimal State Estimation Algorithm

Theorem 1 Based on the descriptions in Sect. 2, suppose we have known the optimal fusion estimation $\hat{x}(k-1|k-1)$ and its estimation error covariance $P(k-1|k-1)$ at time $k-1$, then the optimal state estimation of $x(k)$ at time k could be computed as follows,

$$\hat{x}_{i_j}(k|k) = \hat{x}_{i_{j-1}}(k|k) + K_{i_j}(k)[z_{i_j}(k) - C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k)] \tag{5}$$

$$P_{i_j}(k|k) = P_{i_{j-1}}(k|k) - K_{i_j}(k)[C_{i_j}(k)P_{i_{j-1}}(k|k) + \Delta^T_{i_{j-1}}(k)] \tag{6}$$

$$\begin{aligned} K_{i_j}(k) = {} & [P_{i_{j-1}}(k|k)C^T_{i_j}(k) + \Delta_{i_{j-1}}(k)][C_{i_j}(k)P_{i_{j-1}}(k|k)C^T_{i_j}(k) \\ & + C_{i_j}(k)\Delta_{i_{j-1}}(k) + \Delta^T_{i_{j-1}}(k)C^T_{i_j}(k)]^{-1} + R_{i_j}(k) \end{aligned} \tag{7}$$

$$\begin{aligned} \Delta_{i_j}(k) = {} & \prod_{u=j}^{1}[I - K_{i_u}(k)C_{i_u}(k_{i_u})]S_{i_{j+1}}(k) - K_{i_j}(k)R_{i_j,i_{j+1}}(k) \\ & - \sum_{q=2}^{j}\prod_{u=j}^{q}[I - K_{i_u}(k)C_{i_u}(k_{i_u})]K_{i_{q-1}}(k)R_{i_{q-1},i_{j+1}}(k) \end{aligned} \tag{8}$$

where $j = 1, 2, \ldots, p$. For $j = 0$,

$$\hat{x}_{i_0}(k|k) = \hat{x}(k|k-1) = A(k-1)\hat{x}(k-1|k-1) \tag{9}$$

$$\begin{aligned} P_{i_0}(k|k) &= P(k|k-1) \\ &= A(k-1)P(k-1|k-1)A^T(k-1) + Q(k-1) \end{aligned} \tag{10}$$

$$\Delta_{i_0}(k) = S_{i_1}(k) \tag{11}$$

The above $\hat{x}_{i_j}(k|k)$ and $P_{i_j}(k|k)$ denote the state estimation of $x(k)$ and the corresponding estimation error covariance based on observations of sensors $i_1, i_2, \ldots, i_j$ respectively. When $j = p$, we have $\hat{x}_s(k|k) = \hat{x}_{i_p}(k|k)$ and $P_s(k|k) = P_{i_p}(k|k)$, which are the optimal state fusion estimation and the corresponding estimation error covariance, where subscript 's' means the sequential fusion.

In addition, from (5), we can see that the sensor with the highest sampling rate is the first sensor whose sampling period is assumed to be the unit time, so sensor 1 is sensor i_1. That is, $\hat{x}_{i_1}(k|k) = \hat{x}_1(k|k)$, $P_{i_1}(k|k) = P_1(k|k)$.

Proof The theorem derives from gradually use of the projection theorem. For $i = 1, 2, \ldots, N$, denote

$$Z_i(k) = \{z_i(1), z_i(2), \ldots, z_i(k)\} \tag{12}$$

$$Z^i_1(k) = \{z_1(k), z_2(k), \ldots, z_i(k)\} \tag{13}$$

$$\bar{Z}^i_1(k) = \{Z^i_1(l)\}^k_{l=1} \tag{14}$$

where $Z_i(k)$ is the measurements of sensor i up to time k. If sensor i has no measurement at time l, we denote $z_i(l) = 0$, therefore, the above descriptions are meaningful. $Z^i_1(k)$ is the measurement of sensors $1, 2, \ldots, i$ at time k. $\bar{Z}^i_1(k)$ is the measurements of all sensors at time k and before.

In the sequel, we will prove Theorem 1 deductively by applying the projection theorem. Suppose, we have obtained $\hat{x}_{i_{j-1}}(k|k)$ and the corresponding estimation error covariance $P_{i_{j-1}}(k|k)$, next we will show how to get $\hat{x}_{i_j}(k|k)$ and $P_{i_j}(k|k)$.

Applying the projection theorem, we have

$$\begin{aligned}\hat{x}_{i_j}(k|k) &= E\{x(k)|\bar{Z}_1^N(k-1), Z_1^{i_j}(k)\} \\ &= E\{x(k)|\bar{Z}_1^N(k-1), Z_1^{i_{j-1}}(k), z_{i_j}(k)\} \\ &= \hat{x}_{i_{j-1}}(k|k) + cov\{\tilde{x}_{i_{j-1}}(k|k), \tilde{z}_{i_{j-1}}(k)\} \cdot var\{\tilde{z}_{i_j}(k)\}^{-1}\tilde{z}_{i_j}(k)\end{aligned} \tag{15}$$

where $\tilde{z}_{i_j}(k) = z_{i_j}(k) - \hat{z}_{i_j}(k)$, and

$$\begin{aligned}\hat{z}_{i_j}(k) &= E\{z_{i_j}(k)|\bar{Z}_1^N(k-1), Z_1^{i_{j-1}}(k)\} \\ &= E\{C_{i_j}(k)x(k) + v_{i_j}(k)|\bar{Z}_1^N(k-1), Z_1^{i_{j-1}}(k)\} \\ &= C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k)\end{aligned} \tag{16}$$

So

$$\begin{aligned}\tilde{z}_{i_j}(k) &= z_{i_j}(k) - \hat{z}_{i_j}(k) \\ &= z_{i_j}(k) - C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k) \\ &= C_{i_j}(k)x_{i_j}(k) + v_{i_j}(k) - C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k) \\ &= C_{i_j}(k)\tilde{x}_{i_{j-1}}(k|k) + v_{i_j}(k)\end{aligned} \tag{17}$$

Therefore

$$\begin{aligned}cov\{\tilde{x}_{i_{j-1}}(k|k), \tilde{z}_{i_j}(k)\} &= E\{\tilde{x}_{i_{j-1}}(k|k)\tilde{z}_{i_j}^T(k)\} \\ &= E\{\tilde{x}_{i_{j-1}}(k|k)[C_{i_j}(k)\tilde{x}_{i_{j-1}}(k|k) + v_{i_j}(k)]^T\} \\ &= P_{i_{j-1}}(k|k)C_{i_j}^T(k) + \Delta_{i_{j-1}}(k)\end{aligned} \tag{18}$$

and

$$\begin{aligned}var\{\tilde{z}_{i_j}(k)\} &= E\{\tilde{z}_{i_j}(k)\tilde{z}_{i_j}^T(k)\} \\ &= E\{[C_{i_j}(k)\tilde{x}_{i_{j-1}}(k|k) + v_{i_j}(k)] \cdot [C_{i_j}(k)\tilde{x}_{i_{j-1}}(k|k) + v_{i_j}(k)]^T\} \\ &= C_{i_j}(k)P_{i_{j-1}}(k|k)C_{i_j}^T(k) + R_{i_j}(k) + C_{i_j}(k)\Delta_{i_{j-1}}(k) + \Delta_{i_{j-1}}^T(k)C_{i_j}^T(k)\end{aligned} \tag{19}$$

where

$$\Delta_{i_{j-1}}(k) = E\{\tilde{x}_{i_{j-1}}(k|k)v_{i_j}^T(k)\} \tag{20}$$

By use of the inductive assumption, we have

$$\begin{aligned}\tilde{x}_{i_{j-1}}(k|k) &= x(k) - \hat{x}_{i_{j-1}}(k|k) \\ &= x(k) - \hat{x}_{i_{j-2}}(k|k) - K_{i_{j-1}}(k)[z_{i_{j-1}}(k) - C_{i_{j-1}}(k)\hat{x}_{i_{j-2}}(k|k)] \\ &= \tilde{x}_{i_{j-2}}(k|k) - K_{i_{j-1}}(k)[C_{i_{j-1}}(k)x(k) + v_{i_{j-1}}(k) - C_{i_{j-1}}(k)\hat{x}_{i_{j-2}}(k|k)] \\ &= [I - K_{i_{j-1}}(k)C_{i_{j-1}}(k)]\tilde{x}_{i_{j-2}}(k|k) - K_{i_{j-1}}(k)v_{i_{j-1}}(k)\end{aligned} \tag{21}$$

Substitute (21) into (20), and by use of the inductive hypothesis, we have

$$\begin{aligned}\Delta_{i_{j-1}}(k) &= E\{\tilde{x}_{i_{j-1}}(k|k)v_{i_j}^T(k)\}\\ &= [I - K_{i_{j-1}}(k)C_{i_{j-1}}(k)]E\{\tilde{x}_{i_{j-2}}(k|k)v_{i_j}^T(k)\} - K_{i_{j-1}}(k)E\{v_{i_{j-1}}(k)v_{i_j}^T(k)\}\\ &= \prod_{u=j-1}^{1}[I - K_{i_u}(k)C_{i_u}(k)]S_{i_j}(k) - K_{i_{j-1}}(k)R_{i_{j-1},i_j}(k)\\ &\quad - \sum_{q=2}^{j-1}\prod_{u=j-1}^{q}[I - K_{i_u}(k)C_{i_u}(k)]K_{i_{q-1}}(k)R_{i_{q-1},i_j}(k)\end{aligned} \tag{22}$$

Substitute (18), (19) and the second equation of (17) into (15), we have

$$\hat{x}_{i_j}(k|k) = \hat{x}_{i_{j-1}}(k|k) + K_{i_j}(k)[z_{i_j}(k) - C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k)] \tag{23}$$

where

$$\begin{aligned}K_{i_j}(k) &= cov\{\tilde{x}_{i_{j-1}}(k|k), \tilde{z}_{i_{j-1}}(k)\}var\{\tilde{z}_{i_{j-1}}(k)\}\\ &= [P_{i_{j-1}}(k|k)C_{i_j}^T(k) + \Delta_{i_{j-1}}(k)] \cdot [C_{i_j}(k)P_{i_{j-1}}(k|k)C_{i_j}^T(k) + R_{i_j}(k)\\ &\quad + C_{i_j}(k)\Delta_{i_{j-1}}(k) + \Delta_{i_{j-1}}^T(k)C_{i_j}^T(k)]^{-1}\end{aligned} \tag{24}$$

The estimation error covariance should be computed by

$$\begin{aligned}P_{i_j}(k|k) &= E\{\tilde{x}_{i_j}(k|k)\tilde{x}_{i_j}^T(k|k)\}\\ &= E\{[x(k) - \hat{x}_{i_j}(k|k)][x(k) - \hat{x}_{i_j}(k|k)]^T\}\\ &= E\{[(I - K_{i_j}(k)C_{i_j}(k))\tilde{x}_{i_{j-1}}(k|k) - K_{i_j}(k)v_{i_j}(k)]\\ &\quad \cdot [(I - K_{i_j}(k)C_{i_j}(k))\tilde{x}_{i_{j-1}}(k|k) - K_{i_j}(k)v_{i_j}(k)]^T\}\\ &= [I - K_{i_j}(k)C_{i_j}(k)]P_{i_{j-1}}(k|k)[I - K_{i_j}(k)C_{i_j}(k)]^T\\ &\quad + K_{i_j}(k)R_{i_j}(k)K_{i_j}^T(k) - [I - K_{i_j}(k)C_{i_j}(k)]\Delta_{i_{j-1}}(k)K_{i_j}^T(k)\\ &\quad - K_{i_j}(k)\Delta_{i_{j-1}}^T(k)[I - K_{i_j}(k)C_{i_j}(k)]^T\\ &= P_{i_{j-1}}(k|k) - K_{i_j}(k)[C_{i_j}(k)P_{i_{j-1}}(k|k) + \Delta_{i_{j-1}}^T(k)]\end{aligned} \tag{25}$$

where Eq. (24) is used.

Combine (22), (23), (24), and (25), we obtain

$$\hat{x}_{i_j}(k|k) = \hat{x}_{i_{j-1}}(k|k) + K_{i_j}(k)[z_{i_j}(k) - C_{i_j}(k)\hat{x}_{i_{j-1}}(k|k)] \tag{26}$$

$$P_{i_j}(k|k) = P_{i_{j-1}}(k|k) - K_{i_j}(k) \cdot [C_{i_j}(k)P_{i_{j-1}}(k|k) + \Delta_{i_{j-1}}^T(k)] \tag{27}$$

$$\begin{aligned} K_{i_j}(k) = & [P_{i_{j-1}}(k|k)C_{i_j}^T(k) + \Delta_{i_{j-1}}(k)] \\ & \cdot [C_{i_j}(k)P_{i_{j-1}}(k|k)C_{i_j}^T(k) + R_{i_j}(k) + C_{i_j}(k)\Delta_{i_{j-1}}(k) \\ & + \Delta_{i_{j-1}}^T(k)C_{i_j}^T(k)]^{-1} \end{aligned} \tag{28}$$

$$\begin{aligned} \Delta_{i_j}(k) = & \prod_{u=j}^{1}[I - K_{i_u}(k)C_{i_u}(k_{i_u})]S_{i_{j+1}}(k) - K_{i_j}(k)R_{i_j,i_{j+1}}(k) \\ & - \sum_{q=2}^{j}\prod_{u=j}^{q}[I - K_{i_u}(k)C_{i_u}(k_{i_u})]K_{i_{q-1}}(k)R_{i_{q-1},i_{j+1}}(k) \end{aligned} \tag{29}$$

Let $\hat{x}_s(k|k) = \hat{x}_{i_p}(k|k)$ and $P_s(k|k) = P_{i_p}(k|k)$, then we obtain the state estimation of sequential fusion $\hat{x}_s(k|k)$ and $P_s(k|k)$, and the proof is completed.

4 Simulation

To illustrate the effectiveness of the proposed algorithm, a numerical example is provided in this section.

A target is observed by three sensors, which could be described by Eqs. (1) and (2). Sensor 1 has the highest sampling rate S_1, and the sampling rates of sensor 2 and sensor 3 are S_2 and S_3, respectively, which meet $S_1 = 2S_2 = 3S_3$. And

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, C_1 = [\,1 \quad 0\,], C_2 = [\,1 \quad 0\,], C_3 = [\,0 \quad 1\,]$$

Sensor 1 and sensor 2 observe the position, and sensor 3 observes the velocity.

At time k, the correlations of measurement noises covariance are given by

$$\begin{aligned} R_1(k) = cov(v_1(k)) = 0.048, R_2(k) = cov(v_2(k)) = 0.064 \\ R_3(k) = cov(v_3(k)) = 0.064, R_{12}(k) = E[v_1(k)v_2^T(k)] = 0.032 \\ R_{13}(k) = E[v_1(k)v_3^T(k)] = 0.016, R_{23}(k) = E[v_2(k)v_3^T(k)] = 0.016 \end{aligned}$$

So, the measurement noises covariance is

$$R(k) = \begin{bmatrix} 0.048 & 0.032 & 0.016 \\ 0.032 & 0.064 & 0.016 \\ 0.016 & 0.016 & 0.064 \end{bmatrix}.$$

and $Q(k) = cov(w_k) = \begin{bmatrix} 0.02 & 0.01 \\ 0.01 & 0.04 \end{bmatrix}$. The covariances between the system noise and the measurement noises are given by

$$S_1(k) = E\{w(k-1)v_1^T(k)\} = \begin{bmatrix} 0.0050 \\ 0.0025 \end{bmatrix}$$
$$S_2(k) = E\{w(k-1)v_2^T(k)\} = \begin{bmatrix} 0.0050 \\ 0.0025 \end{bmatrix}$$
$$S_3(k) = E\{w(k-1)v_3^T(k)\} = \begin{bmatrix} 0.0025 \\ 0.0100 \end{bmatrix}$$

To derive $\hat{x}_s(k|k)$, in the measurement update step, if k could be divided by 2 but could not be divided by 3, then we will use the observations of sensor 1 and sensor 2. Similarly, if k could be divided by 3 but could not be divided by 2, then the observations of sensor 1 and sensor 3 should be fused to generate the estimate of $x(k)$. However, if k could be divided by both 2 and 3, that is, k is a multiple of 6, the observations of sensor 1, sensor 2, and sensor 3 should be used. Otherwise, we only use the observations of sensor 1.

The initial conditions are $x_0 = \begin{bmatrix} 10 \\ 0.1 \end{bmatrix}$, $P_0 = 3 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

The Monte Carlo simulation results are shown in Figs. 1, 2, 3, and 4. Using only the measurements of sensor 1, "KF" denotes the Kalman filtering when the sensor noises are cross-correlated and are coupled with the previous step system noise, and "NKF" denotes Kalman filtering when the noises are all independent of each other. Using the measurements of all three sensors, "SFKF" denotes the algorithm given in Theorem 1, and "NSFKF" denotes the sequential fusion algorithm when the noises are treated as independent.

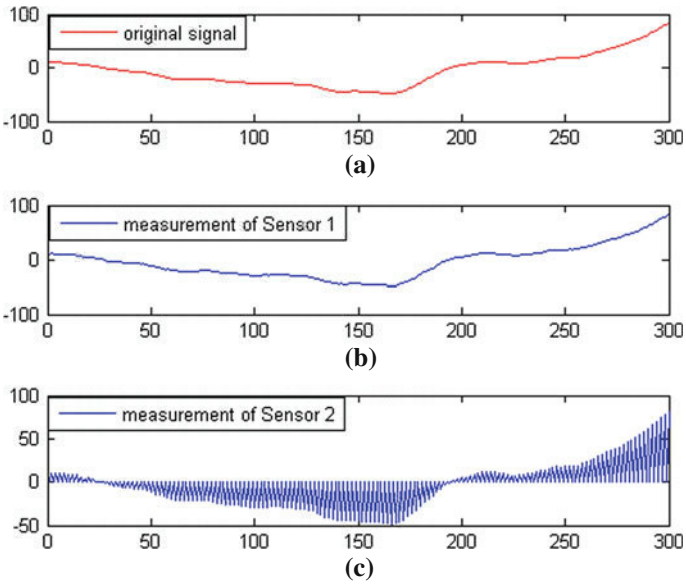


Fig. 1 Position and measurements of sensor 1 and 2

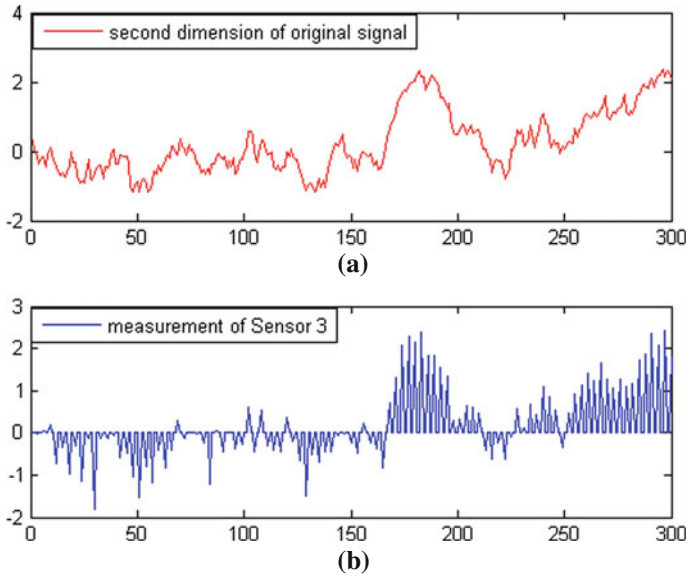


Fig. 2 Velocity and measurements of sensor 3

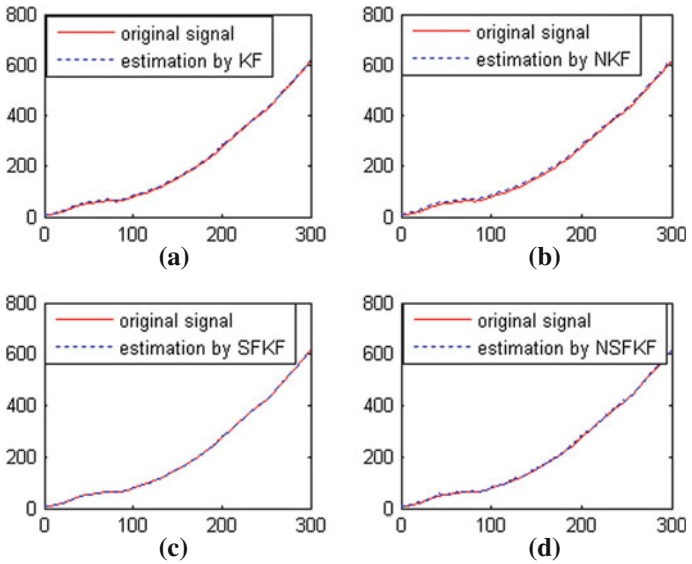


Fig. 3 Position estimations

In Fig. 1, from (a) to (c) are the first dimension of the original signal, the measurement of sensor 1 and the measurements of sensor 2. From the figure, we can see that sensor 2 only has measurements in the even number points and in odd

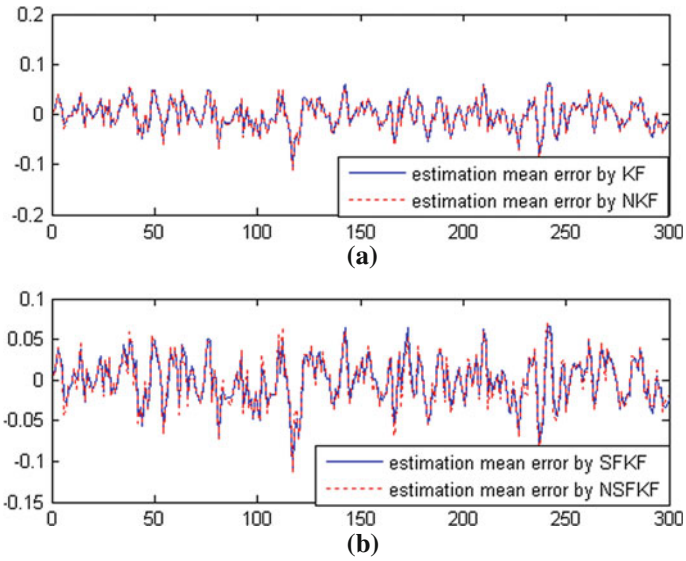


Fig. 4 The statistical position estimation errors of 100 simulations

number points the measurements are zero. In Fig. 2, from (a) to (b) are the second dimension of the original signal and the measurements of sensor 3. Sensor 3 has measurements only when the sampling points are multiple of 3. It can be seen that the measurements are corrupted by noises.

Figure 3 shows the state estimations of the position dimension. From Fig. 3 (a) through (d) are the estimations generated by use of KF, NKF, SFKF, and NSFKF, respectively. For comparison, the estimations are shown in blue dotted line, while the original signal is shown in the red real line. From this figure, it can be seen that all algorithms generate good estimations, whereas the presented algorithm (SFKF) shows the best performance.

In Fig. 4, the statistical estimation errors of 100 simulation runs are shown, where, the position estimation errors of KF and NKF are shown in (a) by the lines of real-blue and dotted-red, respectively. It can be seen from the figure that the errors of KF are slightly less than that of NKF. And (b) shows the position estimation errors of SFKF and NSFKF in blue real lines and red dotted lines, respectively. Also, we can see that the errors of SFKF are less than that of NSFKF.

Briefly, the simulation results in this section illustrate the effectiveness of the presented algorithm.

5 Conclusion

When the sampling rates of different sensors are different and when the measurement noises are cross-correlated and are also coupled with the system noise of the previous step, by use of the projection theory and induction hypothesis repeatedly, a sequential fusion algorithm is generated. The algorithm is proven to be optimal in the sense of Linear Minimum Mean Square Error (LMMSE) mathematically and is applicable to more general cases compared to the existed algorithms.

Acknowledgments The corresponding author of this article is Liping Yan, whose work was supported by the NSFC under grants 61004139 and 91120003, the Scientific research base support, and the outstanding youth foundation of Beijing Institute of Technology. The work of Yuanqing Xia and Mengyin Fu was supported by the NSFC under grants 60974011 and 60904086, respectively. The work of Bo Xiao was supported by Beijing Natural Science Foundation under Grant 4123102, and the innovation youth foundation of Beijing University of Posts and Telecommunications.

References

1. Li XR, Zhu YM, Wang J, Han C (2003) Optimal linear estimation fusion-Part 1:unified fusion rules. IEEE Trans Inf Theory 49:2192–2208
2. Bar-Shalom Y (1990) Multitarget-multisensor tracking: advanced applications, vol 1. Artech House, Norwood
3. Chong CY, Chang KC, Mori S (1986) Distributed tracking in distributed sensor networks. In: 1986 American control conference, Seattle, WA, pp 1863–1868
4. Hashemipour HR, Roy S, Laub AJ (1988) Decentralized structures for parallel Kalman filtering. IEEE Trans Autom Control 3(1):88–93
5. Song E, Zhu Y, Zhou J, You Z (2007) Optimal Kalman filtering fusion with cross-correlated sensor noises. Automatica 43:1450–1456
6. Duan Z, Li XR (2008) The optimality of a class of distributed estimation fusion algorithm. IEEE Inf Fusion 16:1–6
7. Xiao CY, Ma J, Sun SL (2011) Design of information fusion filter for a class of multi-sensor asynchronous sampling systems. In: Control and decision conference, pp 1081–1084
8. Yan LP, Zhou DH, Fu MY, Xia YQ (2010) State estimation for asynchronous multirate multisensor dynamic systems with missing measurements. IET Signal Process 4(6):728–739
9. Shi H, Yan L, Liu B, Zhu J (2008) A sequential asynchronous multirate multisensor data fusion algorithm for state estimation. Chin J Electron 17:630–632

Task Based System Load Balancing Approach in Cloud Environments

Fahimeh Ramezani, Jie Lu and Farookh Hussain

Abstract Live virtual machine (VM) migration is a technique for transferring an active VM from one physical host to another without disrupting the VM. This technique has been proposed to reduce the downtime for migrated overload VMs. As VMs migration takes much more times and cost in comparison with tasks migration, this study develops a novel approach to confront with the problem of overload VM and achieving system load balancing, by assigning the arrival task to another similar VM in a cloud environment. In addition, we propose a multi-objective optimization model to migrate these tasks to a new VM host applying multi-objective genetic algorithm (MOGA). In the proposed approach, there is no need to pause VM during migration time. In addition, as contrast to tasks migration, VM live migration takes longer to complete and needs more idle capacity in host physical machine (PM), the proposed approach will significantly reduce time, downtime memory, and cost consumption.

Keywords Cloud computing · Multi-objective genetic algorithm · Virtual machine migration · Task based system load balancing algorithm

F. Ramezani (✉) · J. Lu · F. Hussain
Decision Systems and e-Service Intelligence Laboratory, Faculty of Engineering and IT, School of Software, Centre for QCIS, University of Technology, Sydney, Australia
e-mail: Fahimeh.Ramezani@students.uts.edu.au

J. Lu
e-mail: Jie.Lu@uts.edu.au

F. Hussain
e-mail: Farookh.Hussain@uts.edu.au

F. Sun et al. (eds.), *Knowledge Engineering and Management*,
Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_4,

1 Introduction

Cloud computing provides new business opportunities for both service providers and requestors clients (e.g., organizations, enterprises, and end users), by means of a platform and delivery model for delivering Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). A cloud encloses the IaaS, PaaS, and/or SaaS inside its own virtualized infrastructure, in order to carry out an abstraction from its underlying physical assets. Typically, the virtualization of a service implies the aggregation of several proprietary processes collected in a virtual environment, called Virtual Machine (VM) [1, 2].

Often, clouds are also spread over distributed virtualization infrastructure covering larger geographical areas (example, let us think about Amazon ('Amazon Elastic Compute Cloud [24]'), Azure [25], and RESERVOIR (an European project facing the cloud computing IaaS topic [3]). In addition, the perspective of cloud federation [4, 5], where cloud providers use virtualized infrastructures of other federated clouds, opens toward new scenarios in which more and more types of new services can be supplied. In fact, clouds exploiting distributed virtualization infrastructures are able to provide new types of "Distributed IaaS, PaaS, and SaaS" [2].

Cloud computing platform using virtualization technology for resource management, achieves dynamic balance between the servers. Using online VM migration technology [6] can online achieve the remapping of VMs and physical resources, and dynamic achieve the whole system load balancing [7]. In modern data center (DC) or cloud environment, virtualization is a critical element since using virtualization the resources can be easily consolidated, partitioned, and isolated. In particular, VM migration has been applied for flexible resource allocation or reallocation, by moving VM from one physical machine to another for stronger computation power, larger memory, fast communication capability, or energy savings [8].

Although a significant amount of research has been done to achieve the whole system load balancing ([6–9], etc.), more improvement is still needed as most of these approaches tried to migrate VMs, when they became overloaded. As VMs migration takes much more times and cost in comparison with tasks migration, we believe migrating tasks from overloaded VMs instead of migrating overloaded VMs, will significantly reduce transfer time and total cost. In addition, to migrate VMs, we have to find a new PM which can accommodate the VM being migrated, and we can rarely avoid choosing an idle PM to optimize power consumption. But in task migration, we just need to find another VM which is located on an active PM and has the same features and number of CPU and more capacity just for executing a task.

Considering these facts and lack of resources in this area, we developed a novel Task Based System Load Balancing (TBSLB) approach to achieve system load balancing and confront with the lack of capacity for executing new task in one VM, by assigning the task to another homogeneous VM in cloud environment. In

addition, we proposed an algorithm to solve the problem of migrating these tasks to new VM host which is a multi-objective problem subject to minimizing cost, minimizing execution time, and transferring time. To solve this problem, we applied multi-objective genetic algorithm (MOGA).

The rest of this paper is organized as follows. In Sect. 2 related works about VM migration are described. In Sect. 3 we propose a conceptual model and the algorithm of TBSLB approach for solving the problem of overloaded VMS by optimal tasks migration from overloaded VMs. The MOGA algorithm is described in Sect. 4. Our developed algorithm for solving multi-objective tasks scheduling problem and completing TBSLB algorithm, is described in Sect. 5. The proposed approach is evaluated in Sect. 6. Finally, we present our conclusion and future work in Sect. 7.

2 Related Works for VM Migration

Virtualization has delivered significant benefits for cloud computing by enabling VM migration to improve utilization, balance load, and alleviate hotspots [10]. Several mechanisms have been proposed to migrate a running instance of a VM (a guest operating system) from one physical host to another to optimize cloud utilization.

VM migration is a hot topic of computing system virtualization. Primary migration relies on process suspend and resume. Many systems [11–13] just pause the VM and copy the state data, then resume the VM on the destination host. This forces the migrated application to stop until all the memory states have been transferred to the migration destination where it is resumed. These methods cause the application to become unavailable during the migration process. ZAP [14] could achieve lower downtime of the service by just transferring a process group, but it still uses stop-and-copy strategy. To reduce the migration downtime and move the VM between hosts in local area network without disrupting it, VMotion [15] and Xen [6] utilize precopy migration technique to perform live migration and support seamless process transfer. Based on their works, [16] tried migrating running VM on a wide area network [8].

In precopy migration technique, VMs migrate by precopying the generated run-time memory state files from the original host to the migration destination host. If the rate for such a dirty memory generation is high, it may take a long time to accomplish live migration because a large amount of data needs to be transferred. In extreme cases, when dirty memory generation rate is faster than precopy speed, live migration will fail. Considering this fact, [8] presented the basic precopy model of VM live migration and proposed an optimized algorithm to improve the performance of live migration by limiting the speed of changing memory through controlling the CPU scheduler of the VM monitor [8].

[7] designed an IPv6 live migration framework for VM based on IPv6 network environment [7]. The framework has been used IPv6 VM live migration in

different IPv6 network. They designed a global control engine as a core complete IPv6 live migration for VM, and provided IPv6 cloud computing service for IPv4/IPv6 client. In their approach, during VM migration process, the source VM would continue to offer services, and the source VM is still not stopped, but no longer provides new services until the old service completion then stop the source VM.

Since power is one of the major limiting factors for a DC or for large cluster growth, [9] proposed a runtime VM mapping framework in a cluster or DC to save energy [9]. Their placement module focused on reducing the power consumption. The main point of their approach is how to map VMs onto a small set of PMs without significant system performance degradation. Actually, they tried to turn off the redundant nodes or PMs to save energy, while the remaining active nodes guarantee the system performance. In their GreenMap framework, one probabilistic, heuristic algorithm is designed employing the idea from simulated annealing (SA) optimization, for the optimization problem: mapping VMs onto a set of PMs under the constraint of multi-dimensional resource consumptions.

[17] believe that most of the proposed methods for on-demand resource provisioning and allocation, focused on the optimization of allocating physical resources to their associated virtual resources, and migrating VMs to achieve load balance and increase resource utilization. Unfortunately, these methods require the suspension of the executing cloud computing applications due to the mandatory shutdown of the associated VMs [17]. To overcome this drawback, they proposed a threshold-based dynamic resource allocation scheme for cloud computing that dynamically allocates the VMs among the cloud computing applications based on their load changes. In their proposed method, they determined when migration should be done but they did not specify the details of how the reallocation will occur.

A fundamental shortcoming of the most existing research is that they consider complete VM migration to overcome overload VM and achieve system load balance. To improve previous approaches and reduce time and cost consumption in such situation, we proposed a new TBSLB approach which migrate tasks from overloads VMs instead of whole VM migration, and to decrease power consumption, a set of VMs on active PMs will be chosen as a new tasks' host. In addition, our approach not only eliminates the process suspend and resume which will happen in VM migration, but also omits precopy mechanism and producing dirty memory in live VM migration.

3 A Conceptual Model and Main Algorithm for Task Based System Load Balancing

In this section, we describe proposed TBSLB approach. This approach contains a conceptual model and TBSLB algorithm which are designed to achieve whole system load balancing by migrating tasks from overloaded VMs. In this approach to

decrease energy consumption and costs, we avoid choosing idle PMs or Computer Nodes (CNs) as a new PM host, because if we transfer tasks to an idle PM, we have to turn it on and this action will increase energy consumption and costs [9].

As cloud computing has the advantages of delivering a flexible, high-performance, pay-as-you-go, on-demand offering service over the Internet, common users and scientists can use cloud computing to solve computationally complex problems (complex applications). The complex applications can be divided into two classes. The one is computing intensive, the other is data intensive. For transferring data intensive applications, the scheduling strategy should decrease the data movement which means decreases the transferring time; but for transferring computing intensive tasks, the scheduling strategy should schedule the data to the high-performance computer [18]. In this paper, we consider bandwidth as a variable to minimize the tasks transferring time for data intensive applications. In addition to enhance performance utilization for computing intensive applications, we consider new host PM's properties (memory, hard disk, etc.).

In cloud environment, there are some tasks schedulers that consider task types, priorities, and their dependencies to schedule tasks in optimal way considering their specific VM's resources. In our proposed system, we design a schedulers' blackboard, where all cloud schedulers (which manage VMs on clouds (see Fig. 1)) share their information about VMs, their features, and their tasks. We apply the information of this blackboard to find an appropriate host VM for the task. Furthermore, the criteria of QoS as SLA information are mentioned in this blackboard.

Every VM has already some tasks to execute and they have limited workload. To determine the time of tasks migration from an overloaded VM, we have to determine online remained workload capacity of a VM (VMs workload information), we defined it as:

$$\mathrm{VM}_{\mathrm{rw}} = \mathrm{VM}_{w} - \mathrm{VM}_{\mathrm{et}} \tag{1}$$

where VM_{w} is VM workload, and $\mathrm{VM}_{\mathrm{et}}$ is the number of executing tasks in VM. The VM will be overload and arrival tasks should be migrated to another similar VM to execute, when:

$$\mathrm{VM}_{\mathrm{rw}} \leq 1 \tag{2}$$

Considering all these facts, we propose a novel TBSLB algorithm which prepares another scheduler to transfer tasks from an overhead VM to a new similar and appropriate VM according to following steps:

Step 1 Gathering data and information about VMMs, VMs, PMs, and SLA information, in the global blackboard as ***inputs*** of TBSLB algorithm as follow:

1. VMs tasks information:

   1.1 The number of executing tasks
   1.2 Tasks' execution time
   1.3 Tasks' performance model

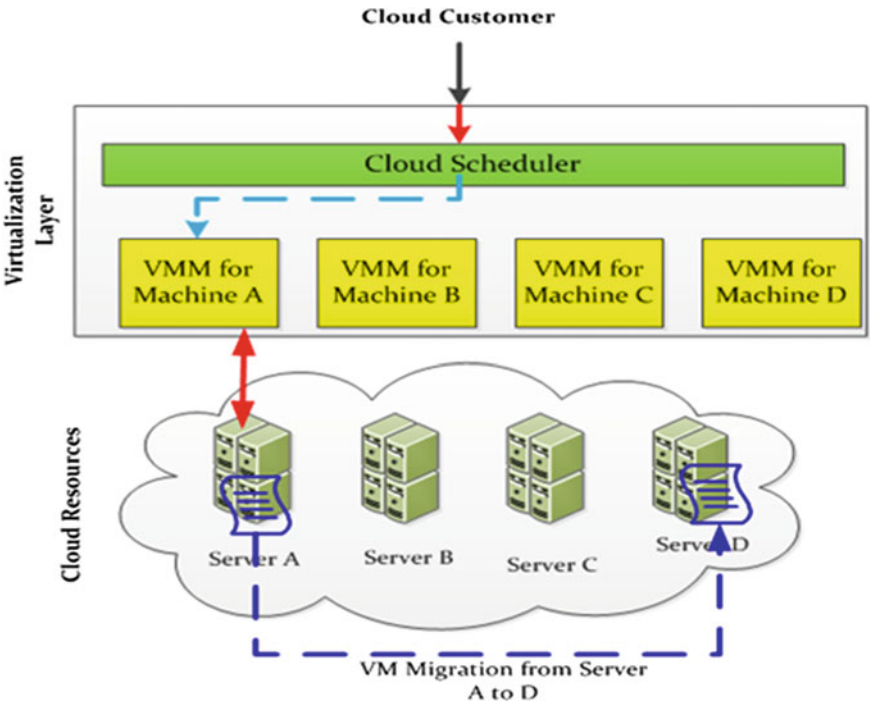


Fig. 1 Cloud architecture

1.4 Tasks' locations
1.5 Tasks' required resources (number of required processors)

2. PMs' Criteria (total/current)

2.1 CPU (number and speed of the processors)
2.2 Free Memory and Hard disk
2.3 Bandwidth
2.4 Idle or active
2.5 Its host VMM

3. SLA information
4. The objectives of the tasks migration optimization model and their information:

4.1 Minimizing cost

4.1.1 Cost information

4.2 Minimizing execution time and transferring time

4.2.1 Execution information
4.2.2 Bandwidth information

Step 2 Monitoring data and information to determine VMs' workflow situation and determining:

1 VMs workload information
2 Overloaded VMs
3 The tasks which should be migrated from overloaded VM's
4 Migration time

Step 3 Finding optimal homogeneous VMs as a new host for executing the tasks of the overloaded VMs, which is a multi-objective task migration problem, applying MOGA (this step will be described in Sect. 5).

Step 4 Considering obtained optimal tasks migration schema, determining following information as the *outputs* of TBSLB algorithm:

1 New optimal cost
2 New optimal execution time
3 Current VMs properties (Executing tasks, CPU, etc.)

Step 5 Transferring tasks and their corresponding data to the optimal host VMs

Step 6 Updating blackboards and schedulers' information according to the outputs of Step 4.

Step 7 End.

The conceptual model of the proposed approach is summarized in Fig. 2.

4 A Multi Objective Genetic Algorithm

In Step 3 of the algorithm, a multi-objective problem which is described in Sect. 5, should be solved to optimize tasks migration from overload VM and find the best VMs as new tasks hosts. In multi-objective optimization problems, each objective function interacts on each other; they almost cannot be optimal at the same time. In other words, one objective function optimization often means a bad developing direction of other objective functions. Therefore, a compromise strategy can be used among the objective functions so as to make them reach optimization at the same time. Now the most popular MOGAs abroad are Corne's PESA2 [19] and PAES [20], SPEA2 [21] proposed by Ziltler, Deb's NSGAII, and so on. Among them, Deb's NSGAII not only has good convergence and distribution but also has higher convergence speed, and solve the shortcomings that shared parameters are difficult to determine [22].

According to MOGA, first initial population whose scale is N is generated randomly. The first generation child population is gained through non-dominated sorting [23] and basic operations such as selection, crossover and mutation. Then, from the second generation on, the parent population and the child population will be merged and sort them based on fast non-dominated. Calculate crowding distance among individuals on each non-dominated layer. According to non-dominant relationship and crowding distance among individuals, select the appropriate

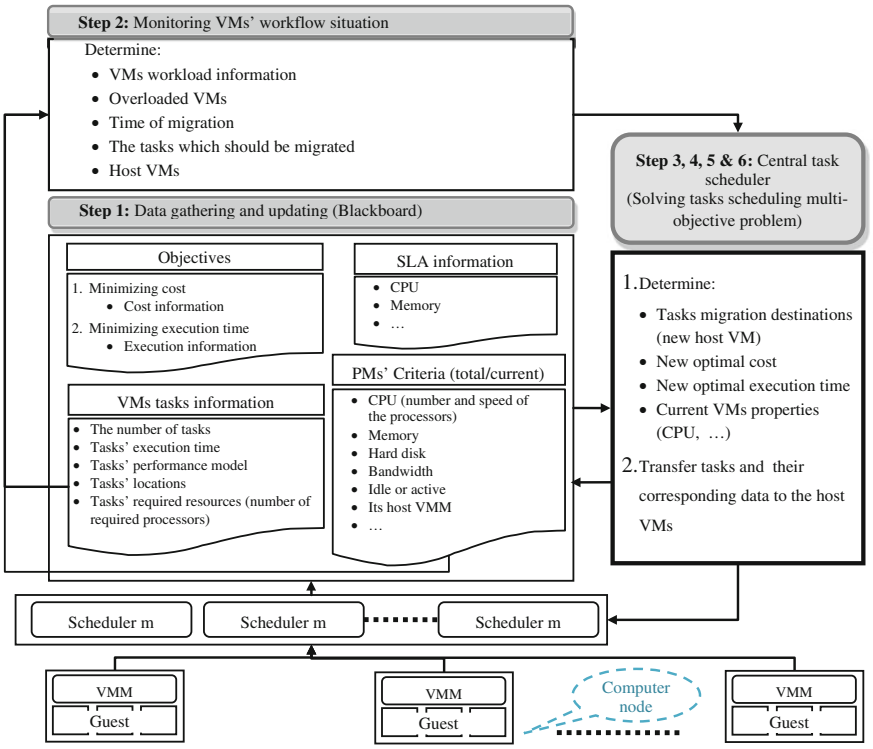


Fig. 2 The conceptual model of TBSLB approach

individuals to form a new parent population. Finally, new child population is generated through basic operations of genetic algorithm. And so on, until the conditions of the process end can be met [22].

5 An Algorithm for Solving Multi-Objective Tasks Migration Problem Using MOGA

In this section, we will describe the sub-TBSLB algorithm which is developed to complete the Step 3 of TBSLB algorithm and solve Multi-objective tasks migration problem. This sub-algorithm will determine the most appropriate VMs to assign the tasks of the overloaded VMs and find optimal tasks scheduling model applying MOGA.

This sub-algorithm applies data and information which are determined in Step 1 and Step 2 of the TBSLB algorithm as its ***inputs***. In this algorithm, we first eliminate those VMs which do not satisfy all constraints to reduce the population

of the candidate solution set. Then we apply MOGA method to find the optimal solution.

According to sub-TBSLB algorithm, to find optimal host VMs to assign new overloaded VMs' tasks, following steps should be conducted:

Step 3.1 Determining candidate host VMs set by choosing the set of VMs which satisfy the constraints about host VMs' properties as $\mathrm{VM}_{\mathrm{set}} = \{\mathrm{vm}_1, \ldots, \mathrm{vm}_m\}$

Step 3.2 Determining overloaded VMs applying Eq. 2, and eliminating them from candidate host VMs set.

Step 3.3 Determining the set of tasks which should migrate from overloaded VMs as immigrating tasks set: $T_{\mathrm{set}} = \{t_1, \ldots, t_n\}$

Step 3.4 Applying MOGA to solve multi-objective problem and assign the immigrating tasks to the optimal host VMs minimizing execution time, transferring time and processing cost. To achieve this goal, following steps should be conducted:

Step 3.4.1 Initializing population P_0 which is generated randomly

Step 3.4.2 Assigning rank to each individual based on non-dominated sort

Step 3.4.3 Implementing binary tournament selection, crossover and mutation on the initial population and creating a new population Q_0 and set $t = 0$

Step 3.4.4 Merging the parent P_t and the child Q_t to form a new population $R_t = P_t \cup Q_t$

Step 3.4.5 Adopting non-dominated relationship to sort population and calculate the crowding distance among population on each layer

Step 3.4.6 Selecting the former N individuals as the parent population, namely $P_{t+1} = P_{t+1}$ [1: N] (Elite strategy)

Step 3.4.7 Implementing reproduction, crossover and mutation on population P_{t+1} to form population Q_{t+1}

Step 3.4.8 If the termination conditions are met, *output* results as optimal tasks migration schema; otherwise, update the evolutionary algebra counter $t = t+1$ and go to step 3.4.4

Step 3.5 End.

6 Evaluation

We determine two parameters to evaluate our proposed TBSLB approach and compare it with traditional whole VM migration methods. The first parameter is related to "power consumption". As, the less number of active PM means the less power consumption [9], we applied following ratio to compare power consumption after load balancing:

$$R_{\mathrm{pc}} = \frac{\text{Number of active PM}}{\text{Number of overloaded VMs}} \tag{3}$$

In proposed approach, to execute some tasks of overloaded VM, we need to find a new similar VM on an active PM as a new host and there will be no need to turn a new PM on. In contrast, for whole VM migration, more hardware capacity will be needed and it is impossible for every case to avoid choosing idle PM. So, in our approach, R_{pc} should be less than whole VM migration attitude for load balancing. Therefore, we will have less "power consumption" after load balancing and:

$$R_{\mathrm{pc}_{\mathrm{Offline\ VM}}} \geq R_{\mathrm{pc}_{\mathrm{Online\ VM}}} > R_{\mathrm{pc}_{\mathrm{NewApproach}}}$$

To compare the efficiency of TBSLB approach, we applied "downtime VM pause time)" as second parameter and estimate the amount of "idle memory" which is prepared during the time of solving the problem of overloaded VM as:

$$M_{\mathrm{im}}(t) = \mathrm{OriginalVM_m}(t) + \mathrm{HostVM_m}(t) \tag{4}$$

where $OriginalVM_m$ and $HostVM_m$ are the amount of original VM memory and host VM, respectively.

In offline VMs migration method, during VM migration time, the original VM should be suspend and its memory and the amount of memory in new host PM which is determined for host VM will be idle. In online VMs migration method, although VM will not be suspended during migration process, the amount of memory in new host PM will be idle in this time. Meanwhile, in TBSLB approach, we eliminate process of suspend and resume in primary VM migration which is mentioned in [8] and there will be no downtime for VMs and no idle memory. As the results:

$$M_{\mathrm{im}}(t)_{\mathrm{Offline\ VM}} > M_{\mathrm{im}}(t)_{\mathrm{Online\ VM}} > M_{\mathrm{im}}(t)_{\mathrm{NewApproach}}$$

7 Conclusion and Future Work

VM migration has been applied for flexible resource allocation or reallocation, by moving overload VM from one PM to another to achieve stronger computation power, larger memory, fast communication capability, or energy savings.

This paper proposed a new TBSLB approach to confront with the problem of overload VM by migrating arrival tasks to another homogeneous VM. This algorithm contains multi-objective tasks migration model subject to minimizing cost, execution time, and transferring time. In proposed approach, there is no need to pause VM during migration time. In addition, as contrast to tasks migration, VM live migration takes longer to complete and needs more idle capacity in host PM, the proposed approach will significantly reduce time, downtime memory, and cost consumption. Furthermore, proposed approach will decrease energy consumption

by avoiding choosing idle PMs or CNs as a new host PM. In our future work, we will propose a method to predict the time of task migration from an overload VM to accelerate load balancing process in our proposed approach.

References

1. Buyya R, Broberg J, Goscinski A (eds) (2011) Cloud computing: principles and paradigms
2. Celesti A, Fazio M, Villari M, Puliafito A (2012) (VM) provisioning through satellite communications in federated cloud environments. Future Gener Comput Syst 28(1):85–93
3. Rochwerger B, Breitgand D, Epstein A, Hadas D, Loy I, Nagin K, Tordsson J, Ragusa C, Villari M, Clayman S (2011) Reservoir-when one cloud is not enough. Comput 44(3):44–51
4. Goiri I, Guitart J, Torres J (2010) Characterizing cloud federation for enhancing providers' profit. In: IEEE 3rd international conference on cloud computing (CLOUD), pp 123–130
5. Ranjan R, Buyya R (2008): Decentralized overlay for federation of enterprise clouds, Arxiv preprint arXiv:0811.2563
6. Clark C, Fraser K, Hand S, Jacob GH (2005) Live migration of (VM)s. In: Proceedings of 2nd ACM/USENIX symposium on network systems, design and implementation (NSDI)
7. Jun C, xiaowei C(2011): IPv6 (VM) live migration framework for cloud computing, Energy procedia, vol 13(0):5753–5757
8. Jin H, Gao W, Wu S, Shi X, Wu X, Zhou F (2011) Optimizing the live migration of (VM) by CPU scheduling'. J of Netw and Comput Appl 34(4):1088–1096
9. Liao X, Jin H, Liu H (2012) Towards a green cluster through dynamic remapping of (VM)s. Future Gener Comput Syst 28(2):469–477
10. Jain N, Menache I, Naor J,.Shepherd F(2012) Topology-aware VM migration in bandwidth oversubscribed datacenter networks, automata, languages, and programming, pp 586–597
11. Kozuch M, Satyanarayanan M (2002) Internet suspend/resume, Mobile computing systems and applications. Proceedings fourth IEEE workshop on, pp 40–46
12. Sapuntzakis CP, Chandra R, Pfaff B, Chow J, Lam MS, Rosenblum M (2002) Optimizing the migration of virtual computers. ACM SIGOPS operating systems review, vol 36, no. SI, pp 377–390
13. Whitaker A, Cox RS, Shaw M, Gribble SD(2004) Constructing services with interposable virtual hardware. In: Proceedings of the 1st symposium on networked systems design and implementation (NSDI), pp 169–182
14. Osman S, Subhraveti D, Su G, Nieh J (2002): The design and implementation of Zap: A system for migrating computing environments, ACM SIGOPS Operating systems review, vol 36, no. SI, pp 361–376
15. Nelson M, Lim BH, Hutchins G (2005) Fast transparent migration for (VM)s, pp 25–25
16. Travostino F, Daspit P, Gommans L, Jog C, De Laat C, Mambretti J, Monga I, Van Oudenaarde B, Raghunath S, Yonghui Wang P (2006) Seamless live migration of (VM)s over the MAN/WAN. Future Gener Comput Syst 22(8):901–907
17. Lin W, Wang JZ, Liang C, Qi D (2011) A threshold-based dynamic resource allocation scheme for cloud computing. Proced Eng 23:695–703
18. Guo L, Zhao S, Shen S, Jiang C (2012) Task scheduling optimization in cloud computing based on heuristic algorithm. J Netw 7(3):547–553
19. Corne DW, Jerram NR, Knowles JD, Oates MJ (2001) PESA-II: Region-based selection in evolutionary multiobjective optimization, Citeseer
20. Knowles JD, Corne DW (2000) Approximating the nondominated front using the Pareto archived evolution strategy. Evolut compu 8(2):149–172
21. Zitzler E, Laumanns M, Thiele L(2001) SPEA2: Improving the strength pareto evolutionary algorithm

22. Zhang Y, Lu C, Zhang H, Han J(2011) Active vibration isolation system integrated optimization based on multi-objective genetic algorithm, computing, control and industrial engineering (CCIE), IEEE 2nd international conference on, vol 1, pp. 258–261
23. Srinivas N, Deb K (1994) Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evol Comput 2(3):221–248
24. Amazon Elastic Compute Cloud (Amazon EC2), http://aws.amazon.com/ec2/
25. Azure: Microsoft's service Cloud platform, http://www.microsoft.com/windowsazure

Accurate Computation of Fingerprint Intrinsic Images with PDE-Based Regularization

Mingyan Li and Xiaoguang Chen

Abstract The intrinsic images of fingerprint, such as orientation field and frequency map, represent the particular and basic characteristics of fingerprint ridge/valley patterns, and play a key role in feature extraction and matching of the fingerprint recognition system. In this paper, a novel algorithm is presented for accurate computation of fingerprint intrinsic images with PDE-based regularization technique. First, the coarse orientation field is estimated using general gradient-based method, and the frequency map is estimated by Fourier spectrum analysis of the projected curve of local window. Then, to measure the reliability of the intrinsic images, the quality map is computed based on the gray-scale intensity and local structural information. Finally, accurate orientation field and frequency map are reconstructed with PDE-based regularization of nonlinear diffusion filtering which is controlled by the quality-based diffusivity. Experimental results illustrated the efficiency of the proposed approach.

Keywords Fingerprint recognition · Orientation field · Frequency map · Image quality measurement · Nonlinear diffusion filtering

1 Introduction

In fingerprint recognition systems, efficient algorithms of fingerprint feature extraction and matching play most important roles [1, 2]. Due to the various reasons, the captured fingerprint images are often degraded and have blurred and

M. Li (✉)
College of Sciences, Beijing Forestry University, 100083 Beijing, China
e-mail: limingyan55@sina.com

X. Chen
Institute of Image Processing and Pattern Recognition, College of Sciences,
North China University of Technology, 100144 Beijing, China
e-mail: chenxg1018@sina.com

F. Sun et al. (eds.), *Knowledge Engineering and Management*,
Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_5, © Springer-Verlag Berlin Heidelberg 2014

Fig. 1 Two sample fingerprint images with fair (*left*) and bad (*right*) quality. The two images are "F0009-08" and "F0022-05" in NIST SD4 [4] respectively

broken ridge structures which decrease the system performance dramatically [1]. To extract reliable features such as minutiae, enhancement algorithms usually are performed to improve the image quality. Most of the existing enhancement algorithms employ some intrinsic characteristics of fingerprint such as ridge orientation and frequency which help to control or adjust the behavior of the designed filters [3].

The most evident structural characteristic of a fingerprint is a pattern of interleaved ridges and valleys, as shown in Fig. 1. The orientation of the ridge pattern flow and the period width of the ridge/valley structure are the critical intrinsic characteristics. For fingerprint images with fair quality, such as the left image in Fig. 1, the intrinsic characteristics can be accurately computed as for most algorithms. However, it remains difficult for low-quality fingerprint images, such as the right image in Fig. 1, because of the heavy noise and serious blurring.

In this paper, we propose a novel approach for computing accurate orientation field and frequency map by utilizing a PDE-based regularization. First, coarse orientation filed is computed by estimating the dominant direction in each local window, and ridge frequency map is computed in Fourier domain by estimating the dominant frequency of the projected curve along the perpendicular direction of the ridge. Then, local quality map of the fingerprint is computed, and it represents the reliability of the computed intrinsic characteristics of the associated local ridge structure. To achieve more accurate estimation, nonlinear diffusion filtering is employed to regularize the coarse orientation and frequency map, and the diffusion process is controlled by the diffusivity which is decided by the local quality score of the fingerprint image. Thus, the accuracy of the intrinsic images is greatly improved after the quality-controlled diffusion process.

The rest of this paper is organized as follows. Section 2 describes the algorithms for computing coarse intrinsic images including orientation field, frequency map, quality map, and segmentation map. In Sect. 3, a PDE-based regularization

algorithm is presented to compute accurate orientation field and frequency map by nonlinear diffusion. Experimental results are presented in Sect. 4. Finally, the paper is concluded with a summary in Sect. 5.

2 Coarse Intrinsic Images Computation

Fingerprint intrinsic images, such as orientation filed and frequency map, represent the essential characteristic of digital image and reflect the main feature of the ridge patterns. In this work, the quality map and segmentation map are also regarded as intrinsic images, since the former indicates the clarity of the ridge patterns and the later indicates the valid region indeed. Therefore, the intrinsic images computed in this paper consist of orientation filed, frequency map, quality map, and segmentation map. It is noted that the mean and variance of the gray-scale intensity of the input fingerprint image are firstly normalized by using the method in [4] before the computation of intrinsic images.

In proposed approach, the input fingerprint image I is divided into nonoverlapped blocks with size $b \times b$. Then the intrinsic characteristics, such as orientation, frequency, quality score, are computed for each block, and the value is assigned to each pixel in the corresponding block. To trade off between computational accuracy and complexity, the block size b is chosen empirically according to the image size. If the image size is small, the value of b even can be set to 1 to perform the computation pixel by pixel.

2.1 Orientation Field

The ridge orientation plays an important role in fingerprint image processing and matching. In the literatures, many approaches have been proposed for estimating the orientation field. In our approach, we adopt the Least-Mean-Square method [3] to coarsely estimate the ridge orientation at each pixel.

For each block centered at pixel (x, y), the orientation of the block is defined as the local dominant ridge orientation of the local window $\mathcal{N}_w(x, y)$ centered at pixel (x, y) with size $w \times w$. The orientation $O(x, y)$ is computed as follows,

$$O(x,y) = \frac{\pi}{2} + \frac{1}{2}\angle\big(S_y(x,y), S_x(x,y)\big), \tag{1}$$

$$S_x(x,y) = \sum_{(i,j)\in\mathcal{N}_w(x,y)} 2I_x(i,j)I_y(i,j), \tag{2}$$

$$S_y(x,y) = \sum_{(i,j)\in\mathcal{N}_w(x,y)} \Big(I_x^2(i,j) - I_y^2(i,j)\Big), \tag{3}$$

where I_x and I_y are the x- and y-component of the image gradient computed by Sobel operators. Here, the size of neighborhood region w is chosen to satisfy $w > b$ which allows a continuous omputation among the adjacent overlapped blocks.

2.2 Frequency Map

The ridge frequency (or period, or width) also is one of the intrinsic characteristics of fingerprint. At the region without minutiae or singular points, the gray intensities along the direction at the alternating ridge and valley forms a curve like sinusoidal wave. Based on this observation, we compute the ridge frequency based on the Fourier spectrum of the projected curve of local oriented window as shown in Fig. 2.

For the block centered at pixel (x, y), a local oriented window with size $L \times w$ centered at pixel (x, y) is defined according to the local ridge orientation as shown in Fig. 2. The pixel intensities in the local window are projected along the ridge orientation to to form a curve. Unlike the method in [4] which is based on the analysis of the peaks of the curve, our approach estimates the ridge frequency based on the Fourier spectrum analysis of the curve. The dominant frequency $F(x, y)$ is calculated by the following formula,

$$F(x,y) = \sum_{k=1}^{L/2} \left(S[k] \frac{k}{L} \right) / \sum_{k=1}^{L/2} S[k], \tag{4}$$

where $S[k]$ is the Fourier spectrum. It is known that the period T of the ridge/valley structure is given by $1/F(x, y)$.

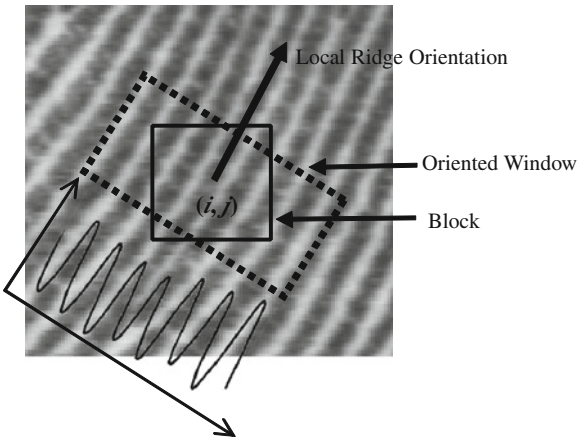


Fig. 2 Ridge frequency estimation based on Fourier spectrum analysis of the projected curve of local oriented window

2.3 Quality Map and Segmentation Map

To measure the accuracy of the orientation and frequency computed in the previous subsections, a quality score for each block is computed based on special features of the local window. Four features are used and there are the range, variance, uniformity of the pixel intensities, and the Fourier spectrum energy.

Range and Variance of Pixel Intensity. The range reflects the difference of the pixel intensity in the window. For the window containing clear ridge/valley pattern, the difference should be large; however, for the noisy or blurred regions, the difference will have smaller range of the pixel intensity. The range is computed as follows,

$$range(x,y) = \overline{max}(I,x,y,w) - \overline{min}(I,x,y,w), \tag{5}$$

where $\overline{max}$ and $\overline{min}$ are the mean value of the αw^2 largest and the αw^2 smallest intensities in the local window $\mathcal{N}_w(x,y)$ respectively. In our experiments, the parameter α is 0.15.

The variance is computed as follows,

$$var(x,y) = \frac{1}{w^2} \sum_{(i,j)\in\mathcal{N}_w(x,y)} (I(i,j) - mean(x,y))^2, \tag{6}$$

where $mean(x,y)$ is the mean value of the pixel intensities in the local window.

Uniformity of Pixel Intensity. The uniformity reflects the distribution on each gray-scale intervals of the pixel intensity . If the block is a normal region containing clear ridge/valley patterns, the pixel intensity should have a uniform distribution. The general gray-scale range [0, 255] is quantized into 32 intervals, and then the statistical histogram is calculated for the window. The uniformity is computed as follows,

$$unif(x,y) = \frac{1}{w^2} \sum_{i=1}^{32} (h[k])^2, \tag{7}$$

where $h[k]$ is the value at the kth bin of the histogram.

Fourier Spectrum Energy. The Fourier spectrum energy reflects the existence of ridge/valley pattern in certain degree, and this feature is computed together with the frequency estimation as described in previous subsection. The computation formula is given by

$$f_{se}(x,y) = \sum_{k=1}^{L/2} c(\|k - k_F\|)s[k], \tag{8}$$

where the coefficient $c(\cdot)$ is Gaussian-like weight function, and $k_F = [F(x,y)/L]$ is the array index corresponding to the dominant frequency.

After the four features are computed as above mentioned, the quality score of the block is computed as follows:

$$Q(x,y) = Q_r(x,y) \times Q_v(x,y) \times Q_u(x,y) \times Q_f(x,y), \tag{9}$$

where Q_r, Q_v, Q_u, Q_f are the quality factors corresponding to the four features respectively, and they are defined as follows,

$$\begin{aligned} Q_r &= 0.5 + 0.5\tanh((range - r_1)/r_2), \\ Q_v &= 0.5 + 0.5\tanh((var - v_1)/v_2), \\ Q_u &= \max\{0, 1 - unif\}, \\ Q_f &= 0.5 + 0.5\tanh((f_{se} - f_1)/f_2), \end{aligned}$$

where parameters $r_1, r_2, v_1, v_2, f_1, f_2$ are chosen empirically.

Based on the quality map Q, the segmentation map R is easily determined by a threshold as follows:

$$R(x,y) = \begin{cases} 1(\text{foreground}) & if\ Q(x,y) > T_Q \\ 0(\text{background}) & \text{otherwise}, \end{cases} \tag{10}$$

where T_Q is the threshold. In our experiment, $T_Q = 0.15$. Post-processing with as morphological operations *close* and *open* are performed to obtain a compact segmentation result.

3 PDE-based Regularization

Due to the continuity of ridge flow, the ridge orientation and frequency also vary continuously and smoothly. However, because of the low quality of the fingerprint image, the computed orientation field and frequency map are not accuracy, and thus the continuity and smoothness are often destroyed. To repair the intrinsic images, post-processing is usually carried out to achieve more accurate ones, and many approaches are proposed in the literatures for this purpose. In [3], low-pass filtering is performed among the adjacent blocks to reduce the noise. In [5], specific filters are designed to estimate the orientation and frequency parameters. Besides, some complicated orientation field models are developed to amend the orientation field based on the singular points, such as in the literatures [6–11]. Considering the effect, the former often fail in the low-quality region, and the latter endures the algorithm complexity and computational burden.

To reconstruct more accurate orientation field and frequency map, a PDE-based regularization approach is proposed which is based on the quality map and non-linear diffusion filtering. In our approach, the local quality score is regarded as the reliability of the computed ridge orientation and frequency, and it determines the diffusivity of the diffusion processing.

Since the orientation value is in $[0, \pi)$, there is ambiguity at the direction 0 and π. To eliminate the ambiguity, the original orientation field O is mapped to a two-dimensional vector field $\Phi = [\Phi_x, \Phi_y]$ by Eq. (11) before regularization, and then regularized Φ will be mapped back to the orientation field O by Eq. (12). Here, Φ is called *Squared Orientation Field (SOF)*.

$$O \mapsto \Phi: \qquad \Phi(x,y) = [\cos(2O(x,y)), \sin(2O(x,y))], \tag{11}$$

$$\Phi \mapsto O: \qquad O(x,y) = \frac{1}{2}\arctan\big(\Phi_y(x,y)/\Phi_x(x,y)\big). \tag{12}$$

For the SOF Φ and frequency map F, we propose to regularize them by the following nonlinear isotropic diffusion,

$$\frac{\partial \Phi(x,y)}{\partial t} = div(C(x,y)\nabla\Phi(x,y)), \tag{13}$$

$$\frac{\partial T(x,y)}{\partial t} = div(C(x,y)\nabla T(x,y)), \;\; T = 1/F, \tag{14}$$

where C is the diffusivity and it control the diffusion strength adaptively.

The local quality score Q indicates the reliability of the ridge information in the region. For low-quality region, stronger diffusion filtering is needed since the information is not reliable; otherwise not much regularization is needed. Therefore, we propose the following formula to map the quality score Q to diffusivity C,

$$C = C_{min} + (1 - C_{min})\left(1 - \tanh\left(\frac{Q - 0.5}{0.15}\right)\right), \tag{15}$$

where C_{min} is the minimum diffusivity to ensure that diffusion filtering will be performed in all valid region. Figure 3 shows the Q–C map function.

As to the numerical implementation of the two PDEs in Eqs. (13) and (14), the semi-implicit scheme with additive operator splitting (AOS) proposed by Weickert et al. [12] is employed in our approach. By using AOS, the computational complexity of solving the PDEs is reduced greatly, and reliable solution can be obtained.

4 Experimental Results

To evaluate the performance, the comparative experiment between proposed algorithm and the algorithm presented in [3] is carried out. A collection of typical fingerprint images from NIST SD4 [4], which covers a variety of different quality levels, is used as benchmark dataset. Two example images are illustrated in Fig. 1.

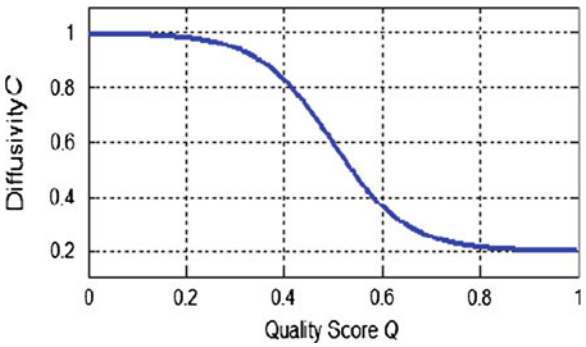


Fig. 3 The map function between quality score Q and diffusivity C. The minimum diffusivity $C_{min} = 0.2$

The results of intrinsic images computation for fingerprint image "F0009-08" in Fig. 1 are depicted in Fig. 4 visually. As seen from the mean/variance-normalized image (Fig. 4a) and estimated quality map (Fig. 4b), the central part has good quality with clear ridge patterns; on the contrary, the upper and lower parts are noisy. What is more, the ridge period in the image is not distributed uniformly, and has a large variety from the lower to the upper parts of the image. As a result, the coarse orientation field (Fig. 4e) and the coarse frequency map (Fig. 4g) do not have sufficient smoothness among different parts. After regularized by diffusion filtering with local quality-based diffusivity (Fig. 4d), continuous and smooth orientation field (Fig. 4f), and frequency map (Fig. 4h) are obtained, which lead to more accurate and robust enhancement processing for feature extraction.

The effectiveness of proposed algorithm is also verified quantitatively by the experiment of minutiae extraction. In this experiment, the intrinsic images of each fingerprint are firstly computed by proposed algorithm and the methods in reference [3], respectively. Then, the Gabor filter-based enhancement algorithm in [3] with different intrinsic images as algorithm parameters is used to improve the quality of fingerprints. Finally, minutiae are extracted from these enhanced fingerprint images. Here, three indicators are computed to describe the performance of minutiae extraction as follows:

$$r_{false} = (n_E - n_M)/n_E, \tag{16}$$

$$r_{miss} = (n_M - n_E)/n_M, \tag{17}$$

$$r_{true} = n_{M \cap E}/n_E, \tag{18}$$

where M and E represent the minutiae sets extracted by human experts and automatic algorithm respectively. n_M, n_E and $n_{M \cap E}$ represent the number of the minutiae in the corresponding sets respectively. By the definitions, r_{false} represents the ratio of spurious minutiae to the extracted minutiae. r_{miss} represents the ratio of missing minutiae to the genuine minutiae. r_{true} represents the ratio of correct

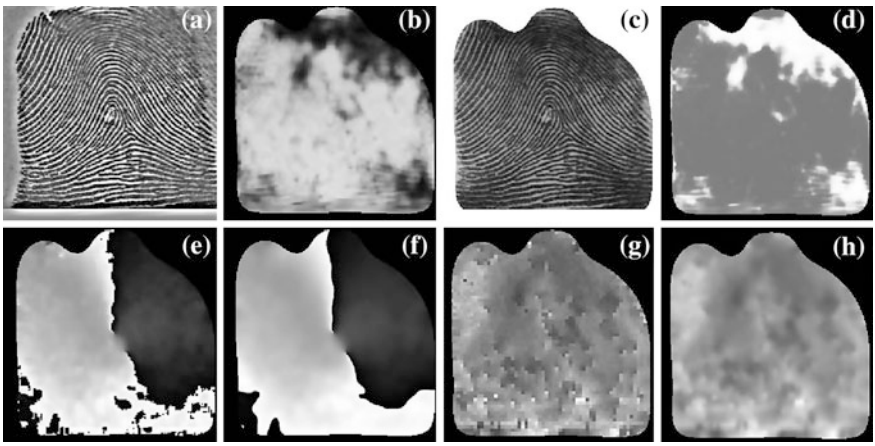


Fig. 4 Intrinsic images computation of the *left* image in Fig. 1. **a** Mean/variance-normalized. **b** Quality map. **c** Segmented fingerprint. **d** Diffusivity. **e** Coarse orientation field. **f** Regularized orientation field. **g** Coarse frequency map. **h** Regularized frequency map. All values displayed in *gray scale*, and frequency maps depict the inverse of the frequency

Table 1 Performance comparison of minutiae extraction

Image	Algorithm in [3]			Proposed algorithm		
	r_{false}	r_{missed}	r_{true}	r_{false}	r_{missed}	r_{true}
F0001_01	0.429	0.075	0.571	0.316	0.075	0.684
F0002_05	0.469	0.254	0.531	0.383	0.186	0.617
F0003_10	0.489	0.274	0.511	0.396	0.294	0.604
F0004_05	0.338	0.122	0.662	0.294	0.102	0.706
F0005_03	0.148	0.137	0.852	0.102	0.041	0.898
F0006_09	0.232	0.059	0.768	0.216	0.029	0.784
F0007_09	0.389	0.121	0.611	0.287	0.060	0.713
F0008_10	0.411	0.166	0.589	0.341	0.083	0.659
F0009_08	0.123	0.040	0.877	0.087	0.013	0.913
F0010_01	0.141	0.106	0.859	0.135	0.146	0.865
Average	0.317	0.135	0.683	**0.256**	**0.103**	**0.744**

minutiae to the extracted minutiae. The larger r_{true} and the smaller r_{false} and r_{miss}, the better the minutiae extraction results. Since the same enhancement algorithm and minutiae extraction algorithm are used, the better performance of minutiae extraction implies the better computation approach of intrinsic images. The performance comparison of minutiae extraction is given in Table 1. From the table, we can observe that the performance of minutiae extraction is substantially improved by using proposed approach.

5 Conclusion

We presented a novel algorithm for accurate computation of fingerprint intrinsic images with PDE-based regularization technique in this paper. First, we estimate the coarse orientation field by gradient-based method and the coarse frequency map by Fourier spectrum analysis of local projected curve. Then, local quality map is computed by a combination of several gray-scale intensity and structural information. Finally, the orientation field and frequency map are adaptively regularized by nonlinear diffusion filtering with quality-based diffusivity. Experimental results show that our proposed approach can compute accurate and satisfactory fingerprint intrinsic images including orientation field, frequency map, quality map and segmentation map, and efficiently improves the performance of enhancement algorithm and minutiae extraction.

Acknowledgments This work was supported by Scientific Research Start-Up Fund of Beijing Forestry University (Grant No. 2010BLX12), Beijing Outstanding Talents Project (No.2011D005002000003), and Beijing Municipal Education Committee Surface Technology Development Plan (No.KM201210009012).

References

1. Maltoni D, Maio D, Jain AK, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
2. Luo X, Tian J, Wu Y (2000) A minutia matching algorithm in fingerprint verification. In: Proceedings of 15th international conference on pattern recognition (ICPR'00), vol 4, pp 833–836
3. Hong L, Wan Y, Jain AK (1998) Fingerprint image enhancement: algorithm and performance evaluation. IEEE Trans Pattern Anal Mach Intell 20(8):777–789
4. NIST fingerprint databases. http://www.nist.gov/srd/biomet.htm
5. Gottschlich C (2012) Curved-region-based ridge frequency estimation and curved Gabor filters for fingerprint image enhancement. IEEE Trans Image Process 21(4):2220–2227
6. Sherlock BG, Monro D (1993) A model for interpreting fingerprint topology. Pattern Recogn 26(7):1047–1055
7. Vizcaya P, Gerhardt L (1996) A nonlinear orientation model for global description of fingerprints. Pattern Recogn 29(7):1221–1231
8. Gu J, Zhou J, Zhang D (2004) A combination model for orientation field of fingerprints. Pattern Recogn 37:543–553
9. Tao X, Yang X, Cao K, Wang R, Li P, Tian J (2010) Estimation of fingerprint orientation field by weighted 2D Fourier expansion model. In: Proceedings of 20th international conference on pattern recognition (ICPR'10), pp 1253–1256
10. Hou Z, Yau W-Y (2010) A variational formulation for fingerprint orientation modeling. In: Proceedings of 20th international conference on pattern recognition (ICPR'10), pp 1626–1629
11. Wang Y, Hu J (2011) Global ridge orientation modeling for partial fingerprint identification. IEEE Trans Pattern Anal Mach Intell 33(1):72–87
12. Weickert J, Romeny B, Viergever M (1998) Effcient and reliable schemes for nonlinear diffusion filtering. IEEE Trans Image Process 7(3):398–411

A Context Ontology Modeling and Uncertain Reasoning Approach Based on Certainty Factor for Context-Aware Computing

Ping Zhang and Hong Huang

Abstract Context-aware computing is an emerging intelligent computing model. Its core idea is to have computing devices understand the real world and automatically provide appropriate services without much concern from users. The research on context modeling and reasoning approaches is the important content of realizing context-aware computing. In this paper, an ontology modeling approach based on certainty factor is proposed in order to not only share and structurally represent context information but also to support the uncertain reasoning by the uncertainty representation for the domain knowledge. Then a certainty factor reasoning approach based on weights is presented to deal with the uncertain context and at last we give a case on the home health telemonitoring application to show the feasibility and validity of the proposed approach.

Keywords Context information · Context modeling · Ontology · Certainty factor · Uncertain reasoning

1 Introduction

With the rapid developments of sensors, computer science, and communication technologies, the intelligence degree of calculation also becomes higher and higher. As a kind of intelligent computing mode, context-aware computing has begun to be recognized by the industry and attracted widespread concerns. It

P. Zhang (✉) · H. Huang
School of Automation, Beijing Institute of Technology, Beijing 100081, People's Republic of China
e-mail: zhangping10003@163.com

H. Huang
e-mail: honghuang@bit.edu.cn

F. Sun et al. (eds.), *Knowledge Engineering and Management*, Advances in Intelligent Systems and Computing 214,
DOI: 10.1007/978-3-642-37832-4_6,

requires that the system should automatically detect the context information (such as users' location, time, environmental parameters, users' activities, and so on), and use these information effectively to compute and provide appropriate services without much concern from users [1]. However, on one hand, various kinds of context information and different ways of perception in the real world make the expression of knowledge and realization of knowledge sharing and reuse become more and more difficult, which will make it hard to implement the knowledge-based reasoning; on the other hand, due to the limitations of sensing devices and network transmission, the context information acquired may be incomplete and imprecise [2]. Meanwhile, because of the inappropriate reasoning method, the high-level context information which is inferred from raw sensing data also has a certain degree of uncertainty [3]. Therefore, for the context-aware computing, how to explicitly describe the context information, realize knowledge sharing and reuse, and at the same time have the capability of modeling and reasoning about the uncertain context are the problems which need to be urgently solved.

The emergence of the ontology model and certainty factor model theory provides a good chance to resolve the problems above. Ontology-based models are more expressive for complex context information than other modeling methods. It becomes possible to share context by providing the formal semantics for context information [4]. But ontology-based models are difficult to support the uncertain reasoning. To overcome the difficulty, we introduce the certainty factor model, which is useful and successful for the uncertain reasoning. Its representative application is MYCIN system [5]. But the certainty factor model is weak in supports for semantic information and it is difficult to achieve the full expression of context information. So in this paper, we put forward an ontology modeling approach based on certainty factor in order to make full use of the complementary advantages of the ontology model and certainty factor model theory. Moreover, the weighting factors are introduced to the model to represent the importance and independence of evidences. On this basis, a certainty factor reasoning approach based on weights is presented to deal with the uncertain context.

This paper is structured as follows. Section 2 outlines the context-aware system. Section 3 gives the research on the ontology model based on certainty factor. Section 4 introduces the certainty factor reasoning approach based on weights. In Sect. 5, we show a simple application case. And Sect. 6 concludes the paper.

2 Overview of the Context-Aware System

Context is any information that can be used to characterize the situation of entities (e.g., a person, place, or object) which are considered relevant to the interaction between a user and an application, including the user and the application themselves [6]. Depending on different properties and purposes, context information can be divided into the following categories:

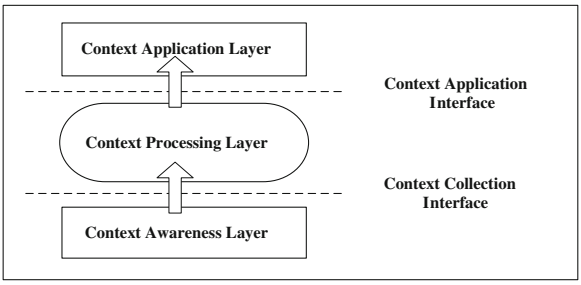


Fig. 1 Hierarchical model of the context-aware system

- User context: user location, personal preferences, social relations.
- Computing context: computing power, network bandwidth, and costs of communication.
- Physical context: temperature, light intensity, noise level.

A system is context-aware if it can acquire, interpret, use context information, and adapt its behavior to the current context of use. As shown in Fig. 1, the context-aware system is logically divided into three layers:

- Context awareness layer: It is mainly used to obtain context information through user inputs or sensors.
- Context processing layer: This layer, which reflects the intelligence of the context-aware system, uses the acquired context information to reason and compute, and get the appropriate processing results.
- Context application layer: According to the context processing results, it will dynamically call the corresponding services, and adjust the device behaviors.

3 Context Model

3.1 Ontology-Based Model

An ontology may be defined as a formal, explicit specification of a shared conceptualization [7]. It helps modeling a world phenomenon by strictly defining its relevant concepts and relationships between the concepts. The ontology-based model has its advantages in: (1) facilitating knowledge sharing by providing a formal specification of the semantics for context information; (2) supporting for logic reasoning, referring to the capability of inferring new context information based on the defined classes and properties; (3) enabling knowledge reuse by use of existing and mature ontology libraries without starting from scratch; (4) having the stronger ability for expressing complex context information.

The Web ontology language (OWL) [8] has become the widely used ontology language for its benefits of the well-defined syntax, efficient reasoning support, formal semantics and the full expression ability, and so on. It can provide more

semantic descriptions than XML, RDF, and RDFS through the increase of some modeling primitives (e.g., owl:Class, owl:DatatypeProperty, and owl:ObjectProperty), greatly improving the expressive ability of the model.

3.2 Ontology Model Based on Certainty Factor

Certainty factor (CF) denotes the degree that people believe that a thing or a phenomenon is true according to historical experiences [9]. It is a subjective and empirical concept and therefore it is difficult to grasp its accuracy. However, for a specific domain, experts have abundant professional knowledge and practical experiences, so it is not difficult to give the certainty factor of domain knowledge. Its value varies from -1 to 1.

The uncertainties of evidences and knowledge are both measured by certainty factor. CF(E) shows the credibility of the evidence E. CF(E) > 0 represents a degree of belief for the evidence E, and CF(E) < 0 represents a degree of disbelief for it. CF(H, E) means the support degree of the evidence E for the conclusion H. Its definition is given as follows [10]:

$$\mathrm{CF}(H,E) = \begin{cases} \frac{P(H/E)-P(H)}{1-P(H)} & P(H/E) \geq P(H) \\ -\frac{P(H)-P(H/E)}{P(H)} & P(H/E) < P(H) \end{cases} \tag{1}$$

According to (1), we can draw that when CF(H, E) > 0, there is $P(H/E)-P(H) > 0$, which means the probability that the conclusion H is true increases because of the presence of the evidence E, and when CF(H, E) $= 0$, there is $P(H/E) = P(H)$, which shows the evidence E has no effects on the conclusion H, and when CF(H, E) < 0, there is $P(H/E)-P(H) < 0$, which indicates the probability that the conclusion H is false increases due to the impacts of the evidence E.

Considering advantages of the ontology in context modeling and supports for uncertain reasoning, this paper proposes an ontology conceptual model based on certainty factor, as is shown in Fig. 2. The context model adopts the hierarchical structure that defines the upper general ontology and the domain-specific ontology. The upper general ontology is composed of abstract concepts of various entities related to context information, in order to achieve ontology sharing in different application environments. The domain-specific ontology defines concept classes closely associated with the specific environments and their concrete subclasses in each sub-domain. The two-tier structure ontology model establishes a loosely coupled relationship between sharing knowledge and specific knowledge, and thus has a certain degree of commonality and expansibility.

In Fig. 2, we have introduced Evidences class and Conclusions class in the upper general ontology, which are used to respectively express the underlying context information obtained from sensors and high-level context information derived by inference. Meanwhile, Certainty Factor class is added to describe the

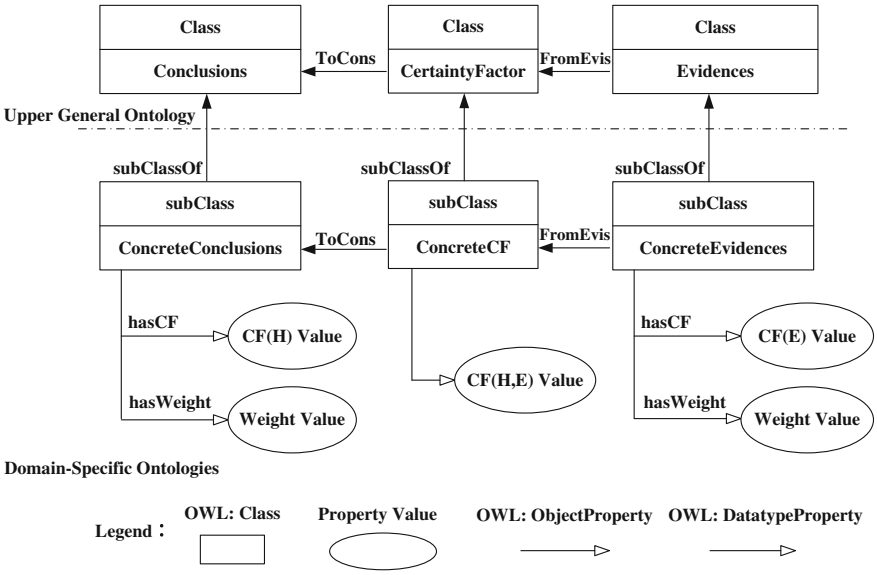


Fig. 2 Ontology conceptual model based on certainty factor

support degree of the evidence for the conclusion. The object properties, FromEvis and ToCons, show relations between the classes, i.e., the interdependencies between conclusions and evidences. In the domain-specific ontology, ConcreteEvidences class, ConcreteConclusions class, and ConcreteCF class, which are subclasses derived through inheritance, are used to denote all possible values involved in a specific domain. The new data type properties, hasCF and hasWeight, are added to these subclasses to express the credibilities and weights of evidences or conclusions, so that the model can support the dynamic nature and uncertainty of context information, and different importance of evidences on conclusions in the context-aware environments. Among them, the weights of evidences are directly given by domain experts. If the conclusion is the final result of reasoning, its weight is set to 1; if the conclusion is the intermediate result of reasoning, it needs to be assigned to the corresponding weight by domain experts in order to continue to serve as the evidence for the upper reasoning.

As the reasoning model of uncertain information, the model not only realizes the precise expression of context information and knowledge sharing, but also supports the uncertain reasoning by the uncertainty representation for the domain knowledge.