

Lecture Notes in Artificial Intelligence 3327

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Yong Shi Weixuan Xu Zhengxin Chen (Eds.)

# Data Mining and Knowledge Management

Chinese Academy of Sciences Symposium  
CASDMKM 2004  
Beijing, China, July 12-14, 2004  
Revised Papers

## Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

## Volume Editors

Yong Shi

Graduate University of Chinese Academy of Sciences  
CAS Research Center on Data Technology and Knowledge Economy  
No. 80 Zhongguancun East Rd., Beijing, China 100080

E-mail: [yshi@mail.unomaha.edu](mailto:yshi@mail.unomaha.edu)

and University of Nebraska at Omaha

College of Information Science and Technology

Omaha, NE 68182, USA

E-mail: [yshi@mail.unomaha.edu](mailto:yshi@mail.unomaha.edu)

Weixuan Xu

Chinese Academy of Sciences, Institute of Policy and Management  
55 Zhongguancun Rd., Beijing 100080, China

E-mail: [wxu@mail.casipm.ac.cn](mailto:wxu@mail.casipm.ac.cn)

Zhengxin Chen

University of Nebraska at Omaha, College of Information Science and Technology  
Omaha, NE 68182, USA

E-mail: [zchen@mail.unomaha.edu](mailto:zchen@mail.unomaha.edu)

Library of Congress Control Number: 2004117657

CR Subject Classification (1998): I.2, H.2.8, H.4, J.1

ISSN 0302-9743

ISBN 3-540-23987-1 Springer Berlin Heidelberg New York

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

Springer is a part of Springer Science+Business Media

[springeronline.com](http://springeronline.com)

© Springer-Verlag Berlin Heidelberg 2004

Printed in Germany

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

Printed on acid-free paper SPIN: 11363613 06/3142 5 4 3 2 1 0

## Preface: Toward an Integrated Study of Data Mining and Knowledge Management

Data mining (DM) and knowledge management (KM) are two important research areas, but with different emphases. Research and practice in these two areas have been largely conducted in parallel. The Chinese Academy of Sciences Symposium on Data Mining and Knowledge Management 2004 (CASDMKM 2004) held in Beijing, China (July 12–14, 2004) provided a unique opportunity for scholars to exchange ideas in these two areas. CASDMKM is a forum for discussing research findings and case studies in data mining, knowledge management and related fields such as machine learning and optimization problems. It promotes data mining technology, knowledge management tools and their real-life applications in the global economy.

This volume of postsymposium proceedings contains 3 invited talks, as well as 25 papers selected from 60 original research papers submitted to the symposium. Contributions in this volume come from scholars within China as well as from abroad, with diverse backgrounds, addressing a wide range of issues. The papers in this volume address various aspects of data mining and knowledge management. We believe the publication of this volume will stimulate the integrated study of these two important areas in the future.

Although both data mining and knowledge management have been active areas in research and practice, there is still a lack of idea exchange between these two camps. CASDMKM aims to bridge this gap. Numerous issues need to be studied in regard to data mining and knowledge management. For example, how to manage the knowledge mined from different data mining methods? From the knowledge management perspective, what kinds of knowledge need to be discovered? What are the similarities and differences for data mining applications and knowledge management applications? What are the issues not yet explored on the boundary of data mining and knowledge management? This list of questions goes on and on. Of course papers in this volume cannot answer all of these questions. Nevertheless, we believe that CASDMKM 2004 served as an exciting platform to foster an integrated study of data mining and knowledge management in the near future.

The papers included in this volume are organized into the following categories:

- *Data mining methods*: Various theoretical aspects of data mining were examined from different perspectives such as fuzzy set theory, linear and non-linear programming, etc.
- *Practical issues of data mining*: Complementary to theoretical studies of data mining, there are also papers exploring aspects of implementing and applying data mining methods.
- *Data mining for bioinformatics*: As a new field, bioinformatics has shown great potential for applications of data mining. The papers included in this category focus on applying data mining methods for microarray data analysis.
- *Data mining applications*: In addition to bioinformatics, data mining methods have also been applied to many other areas. In particular, multiple-

criteria linear and nonlinear programming has proven to be a very useful approach.

- *Knowledge management for enterprise:* These papers address various issues related to the application of knowledge management in corporations using various techniques. A particular emphasis here is on coordination and cooperation.
- *Risk management:* Better knowledge management also requires more advanced techniques for risk management, to identify, control, and minimize the impact of uncertain events, as shown in these papers, using fuzzy set theory and other approaches for better risk management.
- *Integration of data mining and knowledge management:* As indicated earlier, the integration of these two research fields is still in the early stage. Nevertheless, as shown in the papers selected in this volume, researchers have endeavored to integrate data mining methods such as neural networks with various aspects related to knowledge management, such as decision support systems and expert systems, for better knowledge management.

September 2004

Yong Shi  
Weixuan Xu  
Zhengxin Chen

# CASDMKM 2004 Organization

## **Hosted by**

Institute of Policy and Management at the Chinese Academy of Sciences  
Graduate School of the Chinese Academy of Sciences  
International Journal of Information Technology and Decision Making

## **Sponsored by**

Chinese Academy of Sciences  
National Natural Science Foundation of China  
University of Nebraska at Omaha, USA

## **Conference Chairs**

Weixuan Xu, Chinese Academy of Sciences, China  
Yong Shi, University of Nebraska at Omaha, USA

## **Advisory Committee**

Siwei Cheng, Natural Science Foundation, China  
Ruwei Dai, Chinese Academy of Sciences, China  
Masao Fukushima, Kyoto University, Japan  
Bezalel Gavish, Southern Methodist University, USA  
Jiali Ge, Petroleum University, China  
Fred Glover, University of Colorado, USA  
Jifa Gu, Chinese Academy of Sciences, China  
Finn V. Jensen, Aalborg University, Denmark  
Peter Keen, Delft University, Netherlands  
Ralph Keeney, Duke University, USA  
Kin Keung Lai, City University of Hong Kong, Hong Kong, China  
Alexander V. Lotov, Russian Academy of Sciences, Russia  
Robert Nease, Washington University School of Medicine, USA  
Hasan Pirkul, University of Texas at Dallas, USA  
David Poole, University of British Columbia, Canada  
Thomas Saaty, University of Pittsburgh, USA  
Mindia E. Salukvadze, Georgian Academy of Sciences, Georgia  
Elie Sanchez, Université de la Méditerranée, France  
Prakash P. Shenoy, University of Kansas, USA  
Zhongzhi Shi, Chinese Academy of Sciences, China  
Jian Song, Chinese Academy of Engineering, China  
Ralph E. Steuer, University of Georgia, USA  
Peizhuang Wang, Beijing Normal University, China  
Andrew B. Whinston, University of Texas at Austin, USA  
Po-Lung Yu, National Chiao Tung University, Taiwan, and University of Kansas, USA

Philip S. Yu, IBM T.J. Watson Research Center, USA  
Lotfi A. Zadeh, University of California at Berkeley, USA  
Milan Zeleny, Fordham University, USA  
Hans-Jürgen Zimmermann, Aachen Institute of Technology, Germany

### **Program Committee**

Hesham Ali, University of Nebraska at Omaha, USA  
Daobin Chen, Industrial and Commercial Bank of China, China  
Jian Chen, Tsinghua University, China  
Xiaojun Chen, Hirosaki University, Japan  
Zhengxin Chen, University of Nebraska at Omaha, USA  
Chao-Hsien Chu, Pennsylvania State University, USA  
John Chuang, University of California at Berkeley, USA  
Xiaotie Deng, City University of Hong Kong, Hong Kong, China  
Jiawei Han, University of Illinois at Urbana-Champaign, USA  
Xirui Hao, Vision Software Inc., USA  
Chongfu Huang, Beijing Normal University, China  
Haijun Huang, Natural Science Foundation, China  
Zhimin Huang, Adelphi University, USA  
Deepak Khazanchi, University of Nebraska at Omaha, USA  
Wikil Kwak, University of Nebraska at Omaha, USA  
Heeseok Lee, Korea Advanced Institute of Science and Technology, Korea  
Hongyu Li, Fudan University, China  
Shanlyu Li, McGill University, Canada  
Keying Ye, Virginia Polytechnic Institute and State University, USA  
Yachen Lin, First North American Bank, USA  
Jiming Liu, Hong Kong Baptist University, Hong Kong, China  
Xiaohui Liu, Brunel University, UK  
Yoshiteru Nakamori, Japan Advanced Institute of Science and Technology,  
Japan  
David L. Olson, University of Nebraska at Lincoln, USA  
Fuji Ren, Tokushima University, Japan  
Hongchi Shi, University of Missouri-Columbia, USA  
Minghua Shi, Dagong Global Credit Rating Co., China  
Chengzheng Sun, Griffith University, Australia  
Di Sun, China Construction Bank, China  
Minghe Sun, University of Texas at San Antonio, USA  
Tieniu Tan, Chinese Academy of Sciences, China  
Zixiang Tan, Syracuse University, USA  
Xiaowo Tang, Chinese University of Electronic Science and Technology, China  
Xijing Tang, Chinese Academy of Sciences, China  
James Wang, Pennsylvania State University, USA  
Shouyang Wang, Chinese Academy of Sciences, China  
Zhengyuan Wang, University of Nebraska at Omaha, USA  
Yiming Wei, Chinese Academy of Sciences, China  
Xindong Wu, University of Vermont, USA  
Youmin Xi, Xi'an Jiaotong University, China

Lan Xue, Tsinghua University, China

Xiaoguang Yang, Chinese Academy of Sciences, China

Yixian Yang, Beijing University of Posts and Telecommunications, China

Zheng Yang, Sichuan University, China

Gang Yu, University of Texas at Austin, USA

Shuming Zhao, Nanjing University, China

Jianlin Zheng, University of Nebraska Medical Center, USA

Ning Zhong, Maebashi Institute of Technology, Japan

Zongfang Zhou, Chinese University of Electronic Science and Technology, China

Yangyong Zhu, Fudan University, China

# Table of Contents

## Keynote Lectures

Visualization-Based Data Mining Tool and Its Web Application <i>Alexander V. Lotov, Alexander A. Kistanov, Alexander D. Zaitsev</i> .....	1
Knowledge Management, Habitual Domains, and Innovation Dynamics <i>P. L. Yu, T. C. Lai</i> .....	11
Knowledge-Information Circulation Through the Enterprise: Forward to the Roots of Knowledge Management <i>Milan Zeleny</i> .....	22

## Data Mining Methodology

A Hybrid Nonlinear Classifier Based on Generalized Choquet Integrals <i>Zhenyuan Wang, Hai-Feng Guo, Yong Shi, Kwong-Sak Leung</i> .....	34
Fuzzy Classification Using Self-Organizing Map and Learning Vector Quantization <i>Ning Chen</i> .....	41
Solving Discriminant Models Using Interior Point Algorithm <i>Siming Huang, Guoliang Yang, Chao Su</i> .....	51
A Method for Solving Optimization Problem in Continuous Space Using improved Ant Colony Algorithm <i>Ling Chen, Jie Shen, Ling Qin, Jin Fan</i> .....	61

## Practical Issues of Data Mining

Data Set Balancing <i>David L. Olson</i> .....	71
Computation of Least Square Estimates Without Matrix Manipulation <i>Yachen Lin, Chung Chen</i> .....	81
Ensuring Serializability for Mobile Data Mining on Multimedia Objects <i>Shin Parker, Zhengxin Chen, Eugene Sheng</i> .....	90

**Data Mining for Bioinformatics**

“Copasetic Clustering”: Making Sense of Large-Scale Images  
*Karl Fraser, Paul O’Neill, Zidong Wang, Xiaohui Liu*..... 99

Ranking Gene Regulatory Network Models with Microarray Data and Bayesian Network  
*Hongqiang Li, Mi Zhou, and Yan Cui*..... 109

On Efficiency of Experimental Designs for Single Factor cDNA Microarray Experiments  
*Xiao Yang, Keying Ye* ..... 119

**Data Mining Applications**

Data Mining Approach in Scientific Research Organizations Evaluation Via Clustering  
*Jingli Liu, Jianping Li, Weixuan Xu, Yong Shi*..... 128

Heuristics to Scenario-Based Capacity Expansion Problem of PWB Assembly Systems  
*Zhongsheng Hua, Liang Liang*..... 135

A Multiple-Criteria Quadratic Programming Approach to Network Intrusion Detection  
*Gang Kou, Yi Peng, Yong Shi, Zhengxin Chen, Xiaojun Chen*..... 145

Classifications of Credit Cardholder Behavior by Using Multiple Criteria Non-Linear Programming  
*Jing He, Yong Shi, Weixuan Xu*..... 154

Multiple Criteria Linear Programming Data Mining Approach: An Application for Bankruptcy Prediction  
*Wikil Kwak, Yong Shi, John J. Cheh, Heeseok Lee*..... 164

**Knowledge Management for Enterprise**

Coordination and Cooperation in Manufacturer-Retailer Supply Chains  
*Zhimin Huang, Susan X. Li*..... 174

Development of Enterprises’ Capability Based on Cooperative Knowledge Network  
*Junyu Cheng, Hanhui Hu*..... 187

Information Mechanism, Knowledge Management and Arrangement of Corporate Strategem  
*Zhengqing Tang, Jianping Li, Zetao Yan*..... 195

An Integrating Model of Experts' Opinions <i>Jun Tian, Shaochuan Cheng, Kanliang Wang, Yingluo Wang</i> .....	204
<b>Risk Management</b>	
Cartographic Representation of the Uncertainty Related to Natural Disaster Risk: Overview and State of the Art <i>Junxiang Zhang and Chongfu Huang</i> .....	213
A Multi-objective Decision-Making Method for Commercial Banks Loan Portfolio <i>Zhanqin Guo, Zongfang Zhou</i> .....	221
A Multi-factors Evaluation Method on Credit Evaluation of Commerce Banks <i>Zongfang Zhou, Xiaowo Tang, Yong Shi</i> .....	229
<b>Integration of Data Mining and Knowledge Management</b>	
A Novel Hybrid AI System Framework for Crude oil Price Forecasting <i>Shouyang Wang, Lean Yu, K. K. Lai</i> .....	233
A Neural Network and Web-Based Decision Support System for Forex Forecasting and Trading <i>K.K. Lai, Lean Yu, Shouyang Wang</i> .....	243
XML-Based Schemes for Business Project Portfolio Selection <i>Jichang Dong, K. K. Lai, Shouyang Wang</i> .....	254
<b>Author Index</b> .....	263