

LNAI 4632

**Reda Alhajj Hong Gao**  
**Xue Li Jianzhong Li**  
**Osmar R. Zaïane (Eds.)**

# **Advanced Data Mining and Applications**

**Third International Conference, ADMA 2007**  
**Harbin, China, August 2007**  
**Proceedings**



**Springer**

TP274-53  
A193  
2007  
Reda Alhajj Hong Gao Xue Li  
Jianzhong Li Osmar R. Zaiane (Eds.)

# Advanced Data Mining and Applications

Third International Conference, ADMA 2007  
Harbin, China, August 6-8, 2007  
Proceedings



Springer



E2007003093

## Series Editors

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

## Volume Editors

Reda Alhajj  
University of Calgary, Computer Science Department  
Calgary, AB, Canada  
E-mail: alhajj@cpsc.ucalgary.ca

Hong Gao  
Jianzhong Li  
Harbin Institute of Technology, School of Computer Science and Technology  
Harbin, China  
E-mail: {honggao, lijzh}@hit.edu.cn

Xue Li  
The University of Queensland  
School of Information Technology and Electronic Engineering  
Queensland, Australia  
E-mail: xueli@itee.uq.edu.au

Osmar R. Zaiane  
University of Alberta, Department of Computing Science  
Edmonton, AB, Canada  
E-mail: zaiane@cs.ualberta.ca

Library of Congress Control Number: 2007931453

CR Subject Classification (1998): I.2, H.2.8, H.3-4, K.4.4, J.3, I.4, J.1

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN	0302-9743
ISBN-10	3-540-73870-3 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-73870-1 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  
springer.com

© Springer-Verlag Berlin Heidelberg 2007  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper SPIN: 12098095 06/3180 5 4 3 2 1 0

# Lecture Notes in Artificial Intelligence

4632

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

## Preface

The Third International Conference on Advanced Data Mining and Applications (ADMA) organized in Harbin, China continued the tradition already established by the first two ADMA conferences in Wuhan in 2005 and Xi'an in 2006. One major goal of ADMA is to create a respectable identity in the data mining research community. This feat has been partially achieved in a very short time despite the young age of the conference, thanks to the rigorous review process insisted upon, the outstanding list of internationally renowned keynote speakers and the excellent program each year. The impact of a conference is measured by the citations the conference papers receive. Some have used this measure to rank conferences. For example, the independent source [cs-conference-ranking.org](http://cs-conference-ranking.org) ranks ADMA (0.65) higher than PAKDD (0.64) and PKDD (0.62) as of June 2007, which are well established conferences in data mining. While the ranking itself is questionable because the exact procedure is not disclosed, it is nevertheless an encouraging indicator of recognition for a very young conference such as ADMA.

This year we had the pleasure and honour to host illustrious presenters. Our distinguished keynote speakers were Prof. Jaideep Srivastava from Minnesota, who had an enviable career in both industry and academia, and Prof. Geoff Web from Monash University, who is currently the editor-in-chief of the famous journal on Data Mining and Knowledge Discovery published by Springer Netherlands. Our invited speaker is Prof. Zhi-Hua Zhou, a rising star in the Chinese and international research community. Despite his young career, he has rightfully earned global recognition.

ADMA aims at bringing together researchers and practitioners to focus on advancements in data mining and peculiarities and challenges of real world applications using data mining. The major theme of the conference encompasses the innovative applications of data mining approaches to real-world problems that involve large data sets, incomplete and noisy data, or demand optimal solutions. While researchers are eager to submit their novel ideas and techniques covering the first part of the theme of the conference, *advanced data mining*, getting researchers and practitioners to share their experience with applications of data mining involving different knowledge of data in challenging application domains remains limited. This second part of the conference theme, *applications*, is important, as applications are very inspirational and educational, and the organizers continue to encourage the use of ADMA as a venue for sharing such valuable experiences with the data mining community.

This year ADMA received about 200 online submissions from 20 different countries, making it, yet again, a truly international conference. A rigorous process of pre-screening and review involved 92 well known international program committee members and 3 program co-chairs in addition to numerous external reviewers. This screening process yielded the remarkable papers organized in these proceedings in 44 regular papers and 15 short papers, bearing a total acceptance rate of 29%.

We witnessed this year a significant drop in the number of submissions and many authors chose to withdraw their papers after they were accepted. The main reason for

this phenomenon is the fact that Thompson Scientific decided to take Springer's LNCS series out of the ISI journal citation index as of January 2007 arguing that LNCS publications are not journals. In fact, according to their recently published results for 2006, ISI evaluated a total of 25,576 articles published in all 373 journals they classify as computer science journals, whereas LNCS alone published roughly the same number of articles during the same period of time. However, LNCS together with its subseries LNAI and LNBI (like other proceedings from other publishers) will from 2007 onward be included in ISI's newly established Proceedings Index and, of course, will continue to be included in many other well known and frequently used bibliographic indices such as DBLP or CompuServe. Finally, despite the fact that LNCS is no longer covered by ISI/SCI, the standing and recognition of ADMA among the community remains unchanged.

June 2007

Reda Alhajj  
 Hong Gao  
 Xue Li  
 Jianzhong Li  
 Osmar R. Zaiane

# Organization

ADMA 2007 was organized by Harbin Institute of Technology, China.

## Organizers

### General Co-chairs

Jianzhong Li  
Osmar R. Zaiane

Harbin Institute of Technology, China  
University of Alberta, Canada

### Program Co-chairs

Reda Alhajj  
Hong Gao  
Xue Li

University of Calgary, Canada  
Harbin Institute of Technology, China  
Queensland University, Australia

### Local Arrangement Co-chairs

Shengfei Shi  
Lv Tianyang

Harbin Institute of Technology, China  
Harbin Engineering University, China

### Publicity Co-chairs

Haiwei Pan  
Gang Li

Harbin Engineering University, China  
Deakin University, Australia

### Finance Co-chairs

Jizhou Luo  
Randy Goebel

Harbin Institute of Technology, China  
Informatics Circle of Research Excellence, Canada

### Registration Chair

Hongzhi Wang

Harbin Institute of Technology, China

### Web Master

Osmar R. Zaiane

University of Alberta, Canada

## Program Committee

Adam Krzyzak	Concordia University, Montreal, Canada
Ah-Hwee Tan	Nanyang Technological University, Singapore
Alfredo Cuzzocrea	University of Calabria, Italy
Alipio M. Jorge	University of Porto, Portugal
Andre Ponce Leao	University S.Paulo, Brazil
Andrew Kusiak	University of Iowa, USA
Ashkan Sami	Shiraz University, Iran
Arthur Tay	National University of Singapore, Singapore
Bart Goethals	University of Antwerp, Belgium
Carlos Soares	LIACC/Fac. of Economics, University of Porto, Portugal
Carlos Soares	University of Porto, Portugal
Cesar Rego	University of Mississippi, USA
Christophe Giraud-Carrier	Brigham Young University, USA
Daniel C. Neagu	University of Bradford, UK
Deepak S Padmanabhan	IBM India Research Lab, India
Desheng Dash Wu	University of Toronto, Canada
Dimitrios Katsaros	Aristotle University, Greece
Eamonn Keogh	University of California - Riverside, USA
Ee-Peng Lim	Nanyang Technological University, Singapore
Elena Baralis	Politecnico di Torino, Italy
Ezeife Christie	University of Windsor, Canada
Faruk Polat	Middle East Technical University, Turkey
Fernando Berzal	University of Granada, Spain
Francesco Bonchi	KDD Laboratory – ISTI CNR Pisa, Italy
Frans Coenen	The University of Liverpool, UK
Gang Li	Deakin University, Australia
Giovanni Semeraro	University of Bari, Italy
Giuseppe Manco	National Research Council of Italy, Italy
Grigorios Tsoumakas	Aristotle University, Greece
Guoren Wang	NorthEast University, China
Haiwei Pan	Harbin Engineering University, China
Hassan Abolhassani	Sharif University of Technology, Iran
Heng Tao Shen	University of Queensland, Australia
Jaideep Srivastava	University of Minnesota, USA
James Bailey	University of Melbourne, Australia
Jan Rauch	University of Economics, Prague, Czech Republic
Jean-Gabriel G Ganascia	LIP6 - University of Paris, France
Jeffrey Xu Yu	The Chinese University of Hong Kong, Hong Kong, China
Jeremy Besson	Insa-Lyon, France
Jimmy Huang	York University, Canada
Jing Liu	Xidian University, China
JingTao Yao	University of Regina, Canada
Joao Gama	University of Porto, Portugal
Junbin Gao	The University of New England, Australia



Kay Chen Tan	National University of Singapore, Singapore
Krzysztof Cios	University of Colorado at Denver, USA
Longin Jan Latecki	Temple University Philadelphia, USA
Lotfi A. Zadeh	Berkeley University of California, USA
Luis Torgo	University of Porto, Portugal
Mehmed Kantardzic	University of Louisville, USA
Mehmet Kaya	Firat University, Turkey
Michael Frank	University of California, Santa Barbara, USA
Goodchild	
Michael R. Berthold	University of Konstanz, Germany
Mohammad El-Hajj	University of Alberta, Canada
Mohammed Zaki	Rensselaer Polytechnic Institute, USA
Naren Ramakrishnan	Virginia Tech, USA
Nasrullah Memon	Aalborg University, Denmark
Olfa Nasraoui	University of Louisville, USA
Ozgur Ulusoy	Bilkent University, Turkey
Paul Vitanyi	CWI and University of Amsterdam, The Netherlands
Peter Geczy	National Institute of Advanced Industrial Science and Technology, Japan
Philip S. Yu	IBM T.J. Watson Research Center, USA
Ping Jiang	Bradford University, UK
Raul Giraldez Rojo	University of Seville, Spain
Ricardo Vilalta	University of Houston, USA
Rui Camacho	University of Porto, Portugal
Sarah Zelikovitz	College of Staten Island, NY, USA
Shaobin Huang	Harbin Engineering University, China
Shashi Shekhar	University of Minnesota, USA
Shengfei Shi	Harbin Institute of Technology, China
Shichao Zhang	University of Technology, Sydney, Australia
Shuigeng Zhou	Fudan University, China
Shuliang Wang	Wuhan University, China
Shusaku Tsumoto	Shimane Medical University, Japan
Simsek Sule	University of Missouri-Rolla, USA
Sunil Choenni	University of Twente, Netherlands
Tan Kok Kiong	National University of Singapore, Singapore
Tansel Ozyer	TOBB University, Turkey
Vladimir Gorodetsky	Intelligent System Lab, The Russian Academy of Science, Russia
Wanquan Liu	Curtin University of Technology, Australia
Wei Wang	Fudan University, China
Xiangjun Dong	Shandong Institute of Light Industry, China
Xuemin Lin,	University of New South Wales, Australia
Yang ZHANG	Northwest A&F University, China
Yingshu Li	Georgia State University, USA
Yonghong Peng	University of Bradford, UK
Zbigniew W. Ras	University of North Carolina, USA
Zhanhuai Li	Northwest Polytechnical University, China

Zhaoyang Dong	The University of Queensland, Australia
Zhipeng Xie	Fudan University, China
Zijiang Yang	York University, Canada

## **Sponsoring Institutions**

Harbin Institute of Technology  
University of Alberta  
Alberta Innovation and Science  
Harbin Engineering University

# Lecture Notes in Artificial Intelligence (LNAI)

- Vol. 4660: S. Džeroski, J. Todoroski (Eds.), Computational Discovery of Scientific Knowledge. X, 327 pages. 2007.
- Vol. 4632: R. Alhajj, H. Gao, X. Li, J. Li, O.R. Zaiane (Eds.), Advanced Data Mining and Applications. XV, 634 pages. 2007.
- Vol. 4617: V. Torra, Y. Narukawa, Y. Yoshida (Eds.), Modeling Decisions for Artificial Intelligence. XII, 502 pages. 2007.
- Vol. 4612: I. Miguel, W. Ruml (Eds.), Abstraction, Reformulation, and Approximation. XI, 418 pages. 2007.
- Vol. 4604: U. Priss, S. Polovina, R. Hill (Eds.), Conceptual Structures: Knowledge Architectures for Smart Applications. XII, 514 pages. 2007.
- Vol. 4603: F. Pfenning (Ed.), Automated Deduction – CADE-21. XII, 522 pages. 2007.
- Vol. 4597: P. Perner (Ed.), Advances in Data Mining. XI, 353 pages. 2007.
- Vol. 4594: R. Bellazzi, A. Abu-Hanna, J. Hunter (Eds.), Artificial Intelligence in Medicine. XVI, 509 pages. 2007.
- Vol. 4585: M. Kryszkiewicz, J.F. Peters, H. Rybinski, A. Skowron (Eds.), Rough Sets and Intelligent Systems Paradigms. XIX, 836 pages. 2007.
- Vol. 4578: F. Masulli, S. Mitra, G. Pasi (Eds.), Applications of Fuzzy Sets Theory. XVIII, 693 pages. 2007.
- Vol. 4573: M. Kauers, M. Kerber, R. Miner, W. Windsteiger (Eds.), Towards Mechanized Mathematical Assistants. XIII, 407 pages. 2007.
- Vol. 4571: P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. XIV, 913 pages. 2007.
- Vol. 4570: H.G. Okuno, M. Ali (Eds.), New Trends in Applied Artificial Intelligence. XXI, 1194 pages. 2007.
- Vol. 4565: D.D. Schmorow, L.M. Reeves (Eds.), Foundations of Augmented Cognition. XIX, 450 pages. 2007.
- Vol. 4562: D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. XXIII, 879 pages. 2007.
- Vol. 4548: N. Olivetti (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. X, 245 pages. 2007.
- Vol. 4539: N.H. Bshouty, C. Gentile (Eds.), Learning Theory. XII, 634 pages. 2007.
- Vol. 4529: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), Foundations of Fuzzy Logic and Soft Computing. XIX, 830 pages. 2007.
- Vol. 4511: C. Conati, K. McCoy, G. Paliouras (Eds.), User Modeling 2007. XVI, 487 pages. 2007.
- Vol. 4509: Z. Kobti, D. Wu (Eds.), Advances in Artificial Intelligence. XII, 552 pages. 2007.
- Vol. 4496: N.T. Nguyen, A. Grzech, R.J. Howlett, L.C. Jain (Eds.), Agent and Multi-Agent Systems: Technologies and Applications. XXI, 1046 pages. 2007.
- Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), Logic Programming and Nonmonotonic Reasoning. IX, 327 pages. 2007.
- Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. XIV, 585 pages. 2007.
- Vol. 4481: J. Yao, P. Lingras, W.-Z. Wu, M. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), Rough Sets and Knowledge Technology. XIV, 576 pages. 2007.
- Vol. 4476: V. Gorodetsky, C. Zhang, V.A. Skormin, L. Cao (Eds.), Autonomous Intelligent Systems: Multi-Agents and Data Mining. XIII, 323 pages. 2007.
- Vol. 4455: S. Muggleton, R. Otero, A. Tamaddoni-Nezhad (Eds.), Inductive Logic Programming. XII, 456 pages. 2007.
- Vol. 4452: M. Fasli, O. Shehory (Eds.), Agent-Mediated Electronic Commerce. VIII, 249 pages. 2007.
- Vol. 4451: T.S. Huang, A. Nijholt, M. Pantic, A. Pentland (Eds.), Artificial Intelligence for Human Computing. XVI, 359 pages. 2007.
- Vol. 4438: L. Maicher, A. Sigel, L.M. Garshol (Eds.), Leveraging the Semantics of Topic Maps. X, 257 pages. 2007.
- Vol. 4429: R. Lu, J.H. Siekmann, C. Ullrich (Eds.), Cognitive Systems. X, 161 pages. 2007.
- Vol. 4426: Z.-H. Zhou, H. Li, Q. Yang (Eds.), Advances in Knowledge Discovery and Data Mining. XXV, 1161 pages. 2007.
- Vol. 4411: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), Programming Multi-Agent Systems. XIV, 249 pages. 2007.
- Vol. 4410: A. Branco (Ed.), Anaphora: Analysis, Algorithms and Applications. X, 191 pages. 2007.
- Vol. 4399: T. Kovacs, X. Llorà, K. Takadama, P.L. Lanzi, W. Stolzmann, S.W. Wilson (Eds.), Learning Classifier Systems. XII, 345 pages. 2007.
- Vol. 4390: S.O. Kuznetsov, S. Schmidt (Eds.), Formal Concept Analysis. X, 329 pages. 2007.
- Vol. 4389: D. Weyns, H.V.D. Parunak, F. Michel (Eds.), Environments for Multi-Agent Systems III. X, 273 pages. 2007.

- Vol. 4384: T. Washio, K. Satoh, H. Takeda, A. Inokuchi (Eds.), *New Frontiers in Artificial Intelligence. IX*, 401 pages. 2007.
- Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), *Computational Logic in Multi-Agent Systems. X*, 315 pages. 2007.
- Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), *Declarative Programming for Knowledge Management. X*, 229 pages. 2006.
- Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), *Theory and Applications of Relational Structures as Knowledge Instruments II. X*, 373 pages. 2006.
- Vol. 4335: S.A. Brueckner, S. Hassas, M. Jelasity, D. Yamins (Eds.), *Engineering Self-Organising Systems. XII*, 212 pages. 2007.
- Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), *Verification of Object-Oriented Software. XXIX*, 658 pages. 2007.
- Vol. 4333: U. Reimer, D. Karagiannis (Eds.), *Practical Aspects of Knowledge Management. XII*, 338 pages. 2006.
- Vol. 4327: M. Baldoni, U. Endriss (Eds.), *Declarative Agent Languages and Technologies IV. VIII*, 257 pages. 2006.
- Vol. 4314: C. Freksa, M. Kohlhase, K. Schill (Eds.), *KI 2006: Advances in Artificial Intelligence. XII*, 458 pages. 2007.
- Vol. 4304: A. Sattar, B.-h. Kang (Eds.), *AI 2006: Advances in Artificial Intelligence. XXVII*, 1303 pages. 2006.
- Vol. 4303: A. Hoffmann, B.-h. Kang, D. Richards, S. Tsumoto (Eds.), *Advances in Knowledge Acquisition and Management. XI*, 259 pages. 2006.
- Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), *MICA 2006: Advances in Artificial Intelligence. XXVIII*, 1232 pages. 2006.
- Vol. 4289: M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenich, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, M. van Someren (Eds.), *Semantics, Web and Mining. X*, 197 pages. 2006.
- Vol. 4285: Y. Matsumoto, R.W. Sproat, K.-F. Wong, M. Zhang (Eds.), *Computer Processing of Oriental Languages. XVII*, 544 pages. 2006.
- Vol. 4274: Q. Huo, B. Ma, E.-S. Chng, H. Li (Eds.), *Chinese Spoken Language Processing. XXIV*, 805 pages. 2006.
- Vol. 4265: L. Todorovski, N. Lavrač, K.P. Jantke (Eds.), *Discovery Science. XIV*, 384 pages. 2006.
- Vol. 4264: J.L. Balcázar, P.M. Long, F. Stephan (Eds.), *Algorithmic Learning Theory. XIII*, 393 pages. 2006.
- Vol. 4259: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H.S. Nguyen, R. Słowiński (Eds.), *Rough Sets and Current Trends in Computing. XXII*, 951 pages. 2006.
- Vol. 4253: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part III. XXXII*, 1301 pages. 2006.
- Vol. 4252: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXIII*, 1335 pages. 2006.
- Vol. 4251: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXVI*, 1297 pages. 2006.
- Vol. 4248: S. Staab, V. Svátek (Eds.), *Managing Knowledge in a World of Networks. XIV*, 400 pages. 2006.
- Vol. 4246: M. Hermann, A. Voronkov (Eds.), *Logic for Programming, Artificial Intelligence, and Reasoning. XIII*, 588 pages. 2006.
- Vol. 4223: L. Wang, L. Jiao, G. Shi, X. Li, J. Liu (Eds.), *Fuzzy Systems and Knowledge Discovery. XXVIII*, 1335 pages. 2006.
- Vol. 4213: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), *Knowledge Discovery in Databases: PKDD 2006. XXII*, 660 pages. 2006.
- Vol. 4212: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), *Machine Learning: ECML 2006. XXIII*, 851 pages. 2006.
- Vol. 4211: P. Vogt, Y. Sugita, E. Tuci, C.L. Nehaniv (Eds.), *Symbol Grounding and Beyond. VIII*, 237 pages. 2006.
- Vol. 4203: F. Esposito, Z.W. Raś, D. Malerba, G. Semeraro (Eds.), *Foundations of Intelligent Systems. XVIII*, 767 pages. 2006.
- Vol. 4201: Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, E. Tomita (Eds.), *Grammatical Inference: Algorithms and Applications. XII*, 359 pages. 2006.
- Vol. 4200: I.F.C. Smith (Ed.), *Intelligent Computing in Engineering and Architecture. XIII*, 692 pages. 2006.
- Vol. 4198: O. Nasraoui, O. Zaiane, M. Spiliopoulou, B. Mobasher, B. Masand, P.S. Yu (Eds.), *Advances in Web Mining and Web Usage Analysis. IX*, 177 pages. 2006.
- Vol. 4196: K. Fischer, I.J. Timm, E. André, N. Zhong (Eds.), *Multiagent System Technologies. X*, 185 pages. 2006.
- Vol. 4188: P. Sojka, I. Kopeček, K. Pala (Eds.), *Text, Speech and Dialogue. XV*, 721 pages. 2006.
- Vol. 4183: J. Euzenat, J. Domingue (Eds.), *Artificial Intelligence: Methodology, Systems, and Applications. XIII*, 291 pages. 2006.
- Vol. 4180: M. Kohlhase, OMDoc – An Open Markup Format for Mathematical Documents [version 1.2]. XIX, 428 pages. 2006.
- Vol. 4177: R. Marín, E. Onaíndia, A. Bugarín, J. Santos (Eds.), *Current Topics in Artificial Intelligence. XV*, 482 pages. 2006.
- Vol. 4160: M. Fisher, W. van der Hoek, B. Konev, A. Lisitsa (Eds.), *Logics in Artificial Intelligence. XII*, 516 pages. 2006.
- Vol. 4155: O. Stock, M. Schaerf (Eds.), *Reasoning, Action and Interaction in AI Theories and Systems. XVIII*, 343 pages. 2006.
- Vol. 4149: M. Klusch, M. Rovatsos, T.R. Payne (Eds.), *Cooperative Information Agents X. XII*, 477 pages. 2006.

¥ 775.00 元

# Table of Contents

## Invited Talk

Mining Ambiguous Data with Multi-instance Multi-label Representation .....	1
<i>Zhi-Hua Zhou</i>	

## Regular Papers

<i>DELAY: A Lazy Approach for Mining Frequent Patterns over High Speed Data Streams</i> .....	2
<i>Hui Yang, Hongyan Liu, and Jun He</i>	
Exploring Content and Linkage Structures for Searching Relevant Web Pages .....	15
<i>Darren Davis and Eric Jiang</i>	
CLBCRA-Approach for Combination of Content-Based and Link-Based Ranking in Web Search .....	23
<i>Hao-ming Wang and Ye Guo</i>	
Rough Sets in Hybrid Soft Computing Systems .....	35
<i>Renpu Li, Yongsheng Zhao, Fuzeng Zhang, and Lihua Song</i>	
Discovering Novel Multistage Attack Strategies .....	45
<i>Zhitang Li, Aifang Zhang, Dong Li, and Li Wang</i>	
Privacy Preserving DBSCAN Algorithm for Clustering .....	57
<i>K. Anil Kumar and C. Pandu Rangan</i>	
A New Multi-level Algorithm Based on Particle Swarm Optimization for Bisecting Graph .....	69
<i>Lingyu Sun, Ming Leng, and Songnian Yu</i>	
A Supervised Subspace Learning Algorithm: Supervised Neighborhood Preserving Embedding .....	81
<i>Xianhua Zeng and Siwei Luo</i>	
A $k$ -Anonymity Clustering Method for Effective Data Privacy Preservation .....	89
<i>Chuang-Cheng Chiu and Chieh-Yuan Tsai</i>	
LSSVM with Fuzzy Pre-processing Model Based Aero Engine Data Mining Technology .....	100
<i>Xuhui Wang, Shengguo Huang, Li Cao, Dinghao Shi, and Ping Shu</i>	

A Coding Hierarchy Computing Based Clustering Algorithm . . . . .	110
<i>Jing Peng, Chang-jie Tang, Dong-qing Yang, An-long Chen, and Lei Duan</i>	
Mining Both Positive and Negative Association Rules from Frequent and Infrequent Itemsets . . . . .	122
<i>Xiangjun Dong, Zhendong Niu, Xuelin Shi, Xiaodan Zhang, and Donghua Zhu</i>	
Survey of Improving Naive Bayes for Classification . . . . .	134
<i>Liangxiao Jiang, Dianhong Wang, Zhihua Cai, and Xuesong Yan</i>	
Privacy Preserving BIRCH Algorithm for Clustering over Arbitrarily Partitioned Databases . . . . .	146
<i>P. Krishna Prasad and C. Pandu Rangan</i>	
Unsupervised Outlier Detection in Sensor Networks Using Aggregation Tree . . . . .	158
<i>Kejia Zhang, Shengfei Shi, Hong Gao, and Jianzhong Li</i>	
Separator: Sifting Hierarchical Heavy Hitters Accurately from Data Streams . . . . .	170
<i>Yuan Lin and Hongyan Liu</i>	
Spatial Fuzzy Clustering Using Varying Coefficients . . . . .	183
<i>Huaqiang Yuan, Yaxun Wang, Jie Zhang, Wei Tan, Chao Qu, and Wenbin He</i>	
Collaborative Target Classification for Image Recognition in Wireless Sensor Networks . . . . .	191
<i>Xue Wang, Sheng Wang, and Junjie Ma</i>	
Dimensionality Reduction for Mass Spectrometry Data . . . . .	203
<i>Yihui Liu</i>	
The Study of Dynamic Aggregation of Relational Attributes on Relational Data Mining . . . . .	214
<i>Rayner Alfred</i>	
Learning Optimal Kernel from Distance Metric in Twin Kernel Embedding for Dimensionality Reduction and Visualization of Fingerprints . . . . .	227
<i>Yi Guo, Paul W. Kwan, and Junbin Gao</i>	
Efficiently Monitoring Nearest Neighbors to a Moving Object . . . . .	239
<i>Cheqing Jin and Weibin Guo</i>	

A Novel Text Classification Approach Based on Enhanced Association Rule .....	252
<i>Jiangtao Qiu, Changjie Tang, Tao Zeng, Shaojie Qiao, Jie Zuo, Peng Chen, and Jun Zhu</i>	
Applications of the Moving Average of $n^{th}$ -Order Difference Algorithm for Time Series Prediction .....	264
<i>Yang Lan and Daniel Neagu</i>	
Inference of Gene Regulatory Network by Bayesian Network Using Metropolis-Hastings Algorithm .....	276
<i>Khwunta Kirimasthong, Aompilai Manorat, Jeerayut Chaijaruwanich, Sukon Prasitwattanaseree, and Chinnae Thammarongtham</i>	
A Consensus Recommender for Web Users .....	287
<i>Murat Göksedef and Şule Gündüz Ögüdücü</i>	
Constructing Classification Rules Based on SVR and Its Derivative Characteristics .....	300
<i>Dexian Zhang, Zhixiao Yang, Yanfeng Fan, and Ziqiang Wang</i>	
Hiding Sensitive Associative Classification Rule by Data Reduction.....	310
<i>Juggapong Natwichai, Maria E. Orlowska, and Xingzhi Sun</i>	
AOG-ags Algorithms and Applications.....	323
<i>Lizhen Wang, Junli Lu, Joan Lu, and Jim Yip</i>	
A Framework for Titled Document Categorization with Modified Multinomial Naivebayes Classifier .....	335
<i>Hang Guo and Lizhu Zhou</i>	
Prediction of Protein Subcellular Locations by Combining K-Local Hyperplane Distance Nearest Neighbor .....	345
<i>Hong Liu, Haodi Feng, and Daming Zhu</i>	
A Similarity Retrieval Method in Brain Image Sequence Database.....	352
<i>Haiwei Pan, Qilong Han, Xiaoqin Xie, Zhang Wei, and Jianzhong Li</i>	
A Criterion for Learning the Data-Dependent Kernel for Classification .....	365
<i>Jun-Bao Li, Shu-Chuan Chu, and Jeng-Shyang Pan</i>	
Topic Extraction with AGAPE .....	377
<i>Julien Velcin and Jean-Gabriel Ganascia</i>	
Clustering Massive Text Data Streams by Semantic Smoothing Model .....	389
<i>Yubao Liu, Jiarong Cai, Jian Yin, and Ada Wai-Chee Fu</i>	

<i>GraSeq: A Novel Approximate Mining Approach of Sequential Patterns over Data Stream</i> .....	401
<i>Haifeng Li and Hong Chen</i>	
A Novel Greedy Bayesian Network Structure Learning Algorithm for Limited Data .....	412
<i>Feng Liu, Fengzhan Tian, and Qiliang Zhu</i>	
Optimum Neural Network Construction Via Linear Programming Minimum Sphere Set Covering .....	422
<i>Xun-Kai Wei, Ying-Hong Li, and Yu-Fei Li</i>	
How Investigative Data Mining Can Help Intelligence Agencies to Discover Dependence of Nodes in Terrorist Networks .....	430
<i>Nasrullah Memon, David L. Hicks, and Henrik Legind Larsen</i>	
Prediction of Enzyme Class by Using <i>Reactive Motifs</i> Generated from Binding and Catalytic Sites .....	442
<i>Peera Liewlom, Thanawin Rakthanmanon, and Kitsana Waiyamai</i>	
Bayesian Network Structure Ensemble Learning .....	454
<i>Feng Liu, Fengzhan Tian, and Qiliang Zhu</i>	
Fusion of Palmprint and Iris for Personal Authentication .....	466
<i>Xiangqian Wu, David Zhang, Kuanquan Wang, and Ning Qi</i>	
Enhanced Graph Based Genealogical Record Linkage .....	476
<i>Cary Sweet, Tansel Özyer, and Reda Alhajj</i>	
A Fuzzy Comprehensive Clustering Method .....	488
<i>Shuliang Wang and Xinzhou Wang</i>	

**Short Papers**

CACS: A Novel Classification Algorithm Based on Concept Similarity .....	500
<i>Jing Peng, Dong-qing Yang, Chang-jie Tang, Jing Zhang, and Jian-jun Hu</i>	
Data Mining in Tourism Demand Analysis: A Retrospective Analysis ...	508
<i>Rob Law, Henry Mok, and Carey Goh</i>	
Chinese Patent Mining Based on Sememe Statistics and Key-Phrase Extraction .....	516
<i>Bo Jin, Hong-Fei Teng, Yan-Jun Shi, and Fu-Zheng Qu</i>	
Classification of Business Travelers Using SVMs Combined with Kernel Principal Component Analysis .....	524
<i>Xin Xu, Rob Law, and Tao Wu</i>	



Research on the Traffic Matrix Based on Sampling Model . . . . .	533
<i>Fengjun Shang</i>	
A Causal Analysis for the Expenditure Data of Business Travelers . . . . .	545
<i>Rob Law and Gang Li</i>	
A Visual and Interactive Data Exploration Method for Large Data Sets and Clustering . . . . .	553
<i>David Da Costa and Gilles Venturini</i>	
Explorative Data Mining on Stock Data – Experimental Results and Findings . . . . .	562
<i>Lay-Ki Soon and Sang Ho Lee</i>	
Graph Structural Mining in Terrorist Networks . . . . .	570
<i>Muhammad Akram Shaikh, Jiaxin Wang, Zehong Yang, and Yixu Song</i>	
Characterizing Pseudobase and Predicting RNA Secondary Structure with Simple H-Type Pseudoknots Based on Dynamic Programming . . . . .	578
<i>Oyun-Erdene Namsrai and Keun Ho Ryu</i>	
Locally Discriminant Projection with Kernels for Feature Extraction . . . . .	586
<i>Jun-Bao Li, Shu-Chuan Chu, and Jeng-Shyang Pan</i>	
A GA-Based Feature Subset Selection and Parameter Optimization of Support Vector Machine for Content-Based Image Retrieval . . . . .	594
<i>Kwang-Kyu Seo</i>	
E-Stream: Evolution-Based Technique for Stream Clustering . . . . .	605
<i>Komkrit Udommanetanakit, Thanawin Rakthanmanon, and Kitsana Waiyamai</i>	
H-BayesClust: A New Hierarchical Clustering Based on Bayesian Networks . . . . .	616
<i>Morteza Haghir Chehreghani and Hassan Abolhassani</i>	
An Improved AdaBoost Algorithm Based on Adaptive Weight Adjusting . . . . .	625
<i>Lili Cheng, Jianpei Zhang, Jing Yang, and Jun Ma</i>	
<b>Author Index</b> . . . . .	633