Xue Li Osmar R. Zaïane
Zhanhuai Li (Eds.)

Advanced Data Mining and Applications

Second International Conference, ADMA 2006
Xi'an, China, August 14-16, 2006
Proceedings

Springer

Series Editors

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

Volume Editors

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

Osmar R. Zaïane
University of Alberta, Canada
E-mail: zaiane@cs.ualberta.ca

Zhanhuai Li
Northwest Polytechnical University, China
E-mail: lizhh@nwpu.edu.cn

Library of Congress Control Number: 2006930269

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 3-540-37025-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-37025-3 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 2006
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11811305 06/3142 5 4 3 2 1 0

Preface

The Second International Conference on Advanced Data Mining and Applications (ADMA) aimed at establishing its identity in the research community. The theme of ADMA is to focus on the innovative applications of data mining approaches to real-world problems that involve large data sets, incomplete and noisy data, or demand optimal solutions. Data mining is essentially a problem involving different knowledge of data, algorithms, and application domains. The first is about data that are regarded as the "first-class citizens" in application system development. Understanding data is always critical: their structures, high dimensionality, and their qualification and quantification issues. The second is about algorithms: their effectiveness, efficiency, scalability, and their applicability. Amongst a variety of applicable algorithms, selecting a right one to deal with a specific problem is always a challenge that demands contributions from the data mining research community. The third is on domain knowledge of applications. Without a good understanding of domain knowledge, data mining process is hardly able to avoid the GIGO (garbage in and garbage out) effect.

ADMA 2006 received 515 online submissions from 27 countries and regions. A screening process was applied before papers were assigned to reviewers. This eliminated approximately 5% of papers which were not relevant or not worthy for review. The rest of papers were peer reviewed by at least three reviewers consisting of international Program Committee members and external reviewers. We finally accepted 115 papers including 41 regular papers and 74 short papers, yielding a total acceptance rate of 22.3%. For quality control, we conducted a cross-check on the submitted papers for any possible duplicated submissions to four other related concurrent international conferences. Joint actions were taken by the Program Committee Co-chairs from the involved conferences to reject papers that were doubly submitted. In this case, 15 papers were rejected without review crossing PRICAI 2006, ICIC 2006, and ADMA 2006. Two papers were rejected after the acceptance notifications because they appeared in the acceptance list of ICNC-FSKD 2006. While this exercise was very time consuming, it was necessary and we hope that it will send out a message that double submissions are not tolerated in our research community.

The ADMA 2006 program highlights were three keynote speeches from outstanding researchers in advanced data mining and application areas: Usama Fayyad, Jiawei Han, and Ah Chung Tsoi. The conference also invited researchers from two international universities to report on their newest research findings.

August 2006

Xue Li
Osmar Zaïane
Zhanhuai Li

Conference Committee

ADMA 2006 was organized by the University of Queensland, Australia. The major sponsor was Xi'an Software Park, China (*http* : *//www.xasoftpark.com/*). The other sponsor was WISE (Web Information Systems Engineering, *http* : *//www.i − wise.org/*) Society. It was technically co-sponsored by IEEE Queensland Section. The local organization was undertaken by Xi'an Netec Network Technology Co. Ltd. (*http* : *//www.netecweb.com/english/company.asp*).

Organization Committee

Conference Co-chairs	Kyu-Young Whang (Korea Advanced Institute of Science and Technology, Korea)
	Chengqi Zhang (University of Technology, Sydney, Australia)
Program Co-chairs	Xue Li (University of Queensland, Australia)
	Osmar Zaïane (University of Alberta, Canada)
	Zhanhuai Li (Northwest Polytechnical University, China)
Organizing Chairs	Ailiang Mao (Xi'an Software Park, China)
	Jian Chen (Xi'an Software Park–IBM Software Innovation Centre, China)
Publicity Chairs	Yonghong Peng (University of Bradford, UK)
	Jinyan Li (Institute for Infocomm Research, Singapore)
Web Master	Roozbeh Derakhshan (University of Queensland, Australia)

Program Committee Members

Viorel Ariton	Danubius University, Galati, Romania
Michael Bain	University of New South Wales, Australia
Elena Baralis	Politecnico di Torino, Italy
Petr Berka	University of Economics, Czech Republic
Michael R. Berthold	University of Constance, Germany
Fernando Berzal	University of Granada, Spain
Francesco Bonchi	CNR Pisa, Italy
Jean-Francois Boulicaut	INSA Lyon, France
Rui Camacho	University of Porto, Portugal
Xiaochun Cheng	Middlesex University, UK
Krzysztof Cios	University of Colorado, USA

Bruno Cremilleux	Universite de Caen, France
Luc Dehaspe	PharmaDM, Belgium
Zhaoyang Dong	University of Queensland, Australia
Hao Fan	Wuhan University, China
Hongjian Fan	University of Melbourne Australia
Joao Gama	University of Porto, Portugal
Dragan Gamberger	Rudjer Boskovic Institute, Croatia
Jean-Gabriel Ganascia	Universitè Pierre et Marie Curie, France
Junbin Gao	The University of New England, Australia
Christophe Giraud-Carrier	Brigham Young University, USA
Raul Giraldez Rojo	University of Seville, Spain
Bart Goethals	University of Antwerp, Belgium
Vladimir Gorodetsky	The Russian Academy of Science, Russia
Yi Hong	University of South Australia, Australia
Zhanyi Hu	Academy of China, China
Jimmy Huang	York University, Canada
Ping Jiang	Bradford University, UK
Alipio Jorge	University of Porto, Portugal
Mehmed Kantardzic	University of Louisville, USA
Eamonn Keogh	University of California, Riverside, USA
Tim Kovacs	University of Bristol, UK
Adam Krzyzak	Concordia University, Montreal, Canada
Andrew Kusiak	University of Iowa, USA
Longin Jan Latecki	Temple University Philadelphia, USA
Andre Ponce Leao	University of SãoPaulo, Brazil
Gang Li	Deakin University, Australia
Qing Li	Hong Kong City University, Hong Kong, China
Xiaoli Li	Institute for Infocomm Research, Singapore
Xuemin Lin	University of New South Wales, Australia
Wanquan Liu	Curtin University of Technology, Australia
Giuseppe Manco	National Research Council of Italy, Italy
Dunja Mladenic	Jozef Stefan Institute, Slovenia
Iveta Mrazova	Charles University, Prague, Czech Republic
Olfa Nasraoui	University of Louisville, USA
Daniel Neagu	University of Bradford, UK
Claire Nedellec	Laboratoire Mathèmatique, Informatique et Gènome, France
Arlindo Oliveira	Technical University of Lisbon, Portugal
Yonghong Peng	University of Bradford, UK
Jan Rauch	University of Economics, Prague, Czech Republic
Zbigniew W. Ras	University of North Carolina, USA
Cesar Rego	University of Mississippi, USA
Christophe Rigotti	INSA de Lyon, France
Joseph Roure	University of Mataro, Spain
Yucel Saygin	Sabanci University, Turkey
Marc Sebban	University Jean Monnet, France

Giovanni Semeraro	University of Bari, Italy
Seyed A. Shahrestani	University of Western Sydney, Sydney, Australia
Carlos Soares	University of Porto, Portugal
Ah-Hwee Tan	Nanyang Technological University, Singapore
Kay Chen Tan	National University of Singapore, Singapore
Kok Kiong Tan	National University of Singapore, Singapore
Arthur Tay	National University of Singapore, Singapore
Luis Torgo	University of Porto, Portugal
Shusaku Tsumoto	Shimane Medical University, Japan
Brijesh Verma	Central Queensland University, Australia
Ricardo Vilalta	University of Houston, USA
Paul Vitanyi	CWI and University of Amsterdam, The Netherlands
Dianhui Wang	La Trobe University, Melbourne, Australia
Ke Wang	Simon Fraser University, Canada
Shuliang Wang	Wuhan University, China
Wei Wang	University of North Carolina, USA
Hau San Wong	City University of Hong Kong, China
Dash Wu	University of Toronto, Canada
Dongming Xu	University of Queensland, Australia
Qiang Yang	Hong Kong University of Science and Technology, China
Zijiang Yang	York University, Canada
Mao Ye	Electronic Science and Technology, China
Jie Yin	Hong Kong University of Science and Technology, China
Gerson Zaverucha	Federal University of Rio de Janeiro, Brazil
Sarah Zelikovitz	College of Staten Island, NY
Shichao Zhang	University of Technology, Sydney, Australia
Shuigeng Zhou	Fudan University, China
Zhi-Hua Zhou	Nanjin University, China
Zhanli Zhu	Xi'an Shiyou University, China

External Reviewers

Mona Salman Hatam Al-Jiboori	Sultan Qaboos University, Oman
Domingo S. Rodríguez Baena	University of Seville, Spain
Beatriz Pontes Balanza	University of Seville, Spain
Miguel Bugalho	IST/INESC-ID, Portugal
Alexandra Carvalho	IST/INESC-ID, Portugal
Ming Chang	University of Queensland, Australia
Dingyi Chen	University of Queensland, Australia
Luíis Pedro Coelho	IST/INESC-ID, Portugal
Marian Craciun	University "Dunarea de Jos" of Galati, Romania
Roozbeh Derakhshan	University of Queensland, Australia

Yi Ding	University of Queensland, Australia
Xiaoshu Hang	Deakin University, Australia
Petr Hoffmann	Charles University, Prague, Czech Republic
Pavel Jiroutek	Charles University, Prague, Czech Republic
Alexandre Francisco	IST/INESC-ID, Portugal
Ana Teresa Freitas	IST/INESC-ID, Portugal
Gongde Guo	University of Bradford, UK
Yang Lan	University of Bradford, UK
Sara C. Madeira	UBI/INESC-ID/IST, Portugal
Ladan Malazizi	University of Bradford, UK
André Luís Martins	IST/INESC-ID, Portugal
Frantisek Mraz	Charles University, Prague, Czech Republic
Juggapong Natwichai	University of Queensland, Australia
Son Nguyen Nhu	University of Queensland, Australia
Jia Rong	Deakin University, Australia
Luís Silveira Russo	IST/INESC-ID, Portugal
Sule Simsek	University of Missouri-Rolla, USA
Xingzhi Sun	University of Queensland, Australia
Osman Okyar Tahaoglu	Dokuz Eylul University, Turkey
Han (Carol) Tang	University of Missouri-Rolla, USA
Li-Shiang Tsay	Hampton University, USA
Yiqing Tu	Deakin University, Australia
Angelina Tzacheva	University of South Carolina Upstate, USA
Xin Yan	University of Queensland, Australia
Ling Zhuang	Deakin University, Australia

Table of Contents

Invited Papers

Warehousing and Mining Massive RFID Data Sets
Jiawei Han, Hector Gonzalez, Xiaolei Li, Diego Klabjan 1

Self-organising Map Techniques for Graph Data Applications to Clustering of XML Documents
A.C. Tsoi, M. Hagenbuchner, A. Sperduti 19

Finding Time Series Discords Based on Haar Transform
Ada Wai-chee Fu, Oscar Tat-Wing Leung, Eamonn Keogh, Jessica Lin 31

Learning with Local Drift Detection
João Gama, Gladys Castillo 42

Association Rules

A Fast Algorithm for Maintenance of Association Rules in Incremental Databases
Xin Li, Zhi-Hong Deng, Shiwei Tang 56

Extending OLAP with Fuzziness for Effective Mining of Fuzzy Multidimensional Weighted Association Rules
Mehmet Kaya, Reda Alhajj 64

Incremental Maintenance of Association Rules Based on Multiple Previously Mined Results
Zhuohua Duan, Zixing Cai, Yan Lv 72

Mining and Validation of Localized Frequent Web Access Patterns with Dynamic Tolerance
Olfa Nasraoui, Suchandra Goswami 80

SA-IFIM: Incrementally Mining Frequent Itemsets in Update Distorted Databases
Jinlong Wang, Congfu Xu, Hongwei Dan, Yunhe Pan 92

Study of Positive and Negative Association Rules Based on Multi-confidence and Chi-Squared Test
Xiangjun Dong, Fengrong Sun, Xiqing Han, Ruilian Hou 100

Efficiently Mining Maximal Frequent Mutually Associated Patterns
Zhongmei Zhou, Zhaohui Wu, Chunshan Wang, Yi Feng 110

Efficiently Mining Mutually and Positively Correlated Patterns
Zhongmei Zhou, Zhaohui Wu, Chunshan Wang, Yi Feng 118

Classification

ComEnVprs: A Novel Approach for Inducing Decision Tree Classifiers
Shuqin Wang, Jinmao Wei, Junping You, Dayou Liu 126

Towards a Rough Classification of Business Travelers
Rob Law, Thomas Bauer, Karin Weber, Tony Tse 135

Feature Extraction Based on Optimal Discrimination Plane in ECG Signal Classification
Dingfei Ge, Xiao Qu 143

Music Style Classification with a Novel Bayesian Model
Yatong Zhou, Taiyi Zhang, Jiancheng Sun 150

Classification of Polarimetric SAR Data Based on Multidimensional Watershed Clustering
Wen Yang, Hao Wang, Yongfeng Cao, Haijian Zhang 157

An Effective Combination Based on Class-Wise Expertise of Diverse Classifiers for Predictive Toxicology Data Mining
Daniel Neagu, Gongde Guo, Shanshan Wang 165

Robust Collective Classification with Contextual Dependency Network Models
Yonghong Tian, Tiejun Huang, Wen Gao 173

User-Centered Image Semantics Classification
Hongli Xu, De Xu, Fangshi Wang 181

A Performance Study of Gaussian Kernel Classifiers for Data Mining Applications
Miyoung Shin 189

TTLSC – Transductive Total Least Square Model for Classification and Its Application in Medicine
Qun Song, Tian Min Ma, Nikola Kasabov 197

Forecasting Electricity Market Price Spikes Based on Bayesian Expert with Support Vector Machines
Wei Wu, Jianzhong Zhou, Li Mo, Chengjun Zhu 205

Integrating Local One-Class Classifiers for Image Retrieval
Yiqing Tu, Gang Li, Honghua Dai 213

Incremental Discretization for Naïve-Bayes Classifier
Jingli Lu, Ying Yang, Geoffrey I. Webb 223

Distance Guided Classification with Gene Expression Programming
Lei Duan, Changjie Tang, Tianqing Zhang, Dagang Wei, Huan Zhang 239

Research on Multi-valued and Multi-labeled Decision Trees
Hong Li, Rui Zhao, Jianer Chen, Yao Xiang 247

Clustering

A Spatial Clustering Algorithm Based on SOFM
Zhong Qu, Lian Wang 255

Mining Spatial-temporal Clusters from Geo-databases
Min Wang, Aiping Wang, Anbo Li 263

A Fuzzy Subspace Algorithm for Clustering High Dimensional Data
Guojun Gan, Jianhong Wu, Zijiang Yang 271

Robust Music Information Retrieval on Mobile Network Based on Multi-Feature Clustering
Won-Jung Yoon, Sanghun Oh, Kyu-Sik Park 279

Joint Cluster Based Co-clustering for Clustering Ensembles
Tianming Hu, Liping Liu, Chao Qu, Sam Yuan Sung 284

Mining Gait Pattern for Clinical Locomotion Diagnosis Based on Clustering Techniques
Guandong Xu, Yanchun Zhang, Rezaul Begg 296

Combining Multiple Clusterings Via k-Modes Algorithm
Huilan Luo, Fansheng Kong, Yixiao Li 308

HOV^3: An Approach to Visual Cluster Analysis
Ke-Bing Zhang, Mehmet A. Orgun, Kang Zhang 316

A New Fuzzy Co-clustering Algorithm for Categorization of Datasets with Overlapping Clusters
William-Chandra Tjhi, Lihui Chen 328

Quantum-Behaved Particle Swarm Optimization Clustering Algorithm
Jun Sun, Wenbo Xu, Bin Ye 340

Clustering Mixed Data Based on Evidence Accumulation
Huilan Luo, Fansheng Kong, Yixiao Li 348

Mining Maximal Local Conserved Gene Clusters from Microarray Data
Yuhai Zhao, Guoren Wang, Ying Yin, Guangyu Xu 356

Novel Algorithms

A Novel P2P Information Clustering and Retrieval Mechanism
Huaxiang Zhang, Peide Liu 364

Keeping Track of Customer Life Cycle to Build Customer Relationship
Sung Ho Ha, Sung Min Bae 372

Mining of Flexible Manufacturing System Using Work Event Logs and Petri Nets
Hesuan Hu, Zhiwu Li, Anrong Wang 380

Improved Genetic Algorithm for Multiple Sequence Alignment Using Segment Profiles (GASP)
Yanping Lv, Shaozi Li, Changle Zhou, Wenzhong Guo, Zhengming Xu 388

A Novel Visual Clustering Algorithm for Finding Community in Complex Network
Shuzhong Yang, Siwei Luo, Jianyu Li 396

Self-Organizing Network Evolving Model for Mining Network Community Structure
Bo Yang 404

An Interactive Visualization Environment for Data Exploration Using Points of Interest
David Da Costa, Gilles Venturini 416

Forecasting the Volatility of Stock Price Index
Tae Hyup Roh 424

ExMiner: An Efficient Algorithm for Mining Top-K Frequent Patterns
Tran Minh Quang, Shigeru Oyanagi, Katsuhiro Yamazaki 436

Learning Bayesian Networks Structure with Continuous Variables
Shuang-Cheng Wang, Xiao-Lin Li, Hai-Yan Tang 448

A Unified Strategy of Feature Selection
Peng Liu, Naijun Wu, Jiaxian Zhu, Junjie Yin, Wei Zhang 457

Experimental Comparison of Feature Subset Selection Using GA and ACO Algorithm
Keunjoon Lee, Jinu Joo, Jihoon Yang, Vasant Honavar 465

OMVD: An Optimization of MVD
Zhi He, Shengfeng Tian, Houkuan Huang 473

ZED: Explaining Temporal Variations in Query Volume
Maojin Jiang, Shlomo Argamon, Abdur Chowdhury, Kush Sidhu 485

An Effective Multi-level Algorithm for Bisecting Graph
Ming Leng, Songnian Yu 493

A New Polynomial Time Algorithm for Bayesian Network Structure Learning
Sanghack Lee, Jihoon Yang, Sungyong Park 501

Personalized Recommendation Based on Partial Similarity of Interests
Ming-Hua Yang, Zhi-Min Gu 509

A Fast Implementation of the EM Algorithm for Mixture of Multinomials
Jan Peter Patist 517

A Novel Approach to Pattern Recognition Based on PCA-ANN in Spectroscopy
Xiaoli Li, Yong He 525

Semi-supervised Dynamic Counter Propagation Network
Yao Chen, Yuntao Qian 533

The Practical Method of Fractal Dimensionality Reduction Based on Z-Ordering Technique
Guanghui Yan, Zhanhuai Li, Liu Yuan 542

Feature Selection for Complex Patterns
Peter Schenkel, Wanqing Li, Wanquan Liu 550

Naïve Bayesian Tree Pruning by Local Accuracy Estimation
Zhipeng Xie 558

A New Visualization Method for Patent Map: Application to Ubiquitous Computing Technology
Jong Hwan Suh, Sang Chan Park 566

Local Linear Logistic Discriminant Analysis with Partial Least Square Components
Jangsun Baek, Young Sook Son 574

Activity Mining: Challenges and Prospects
Longbing Cao 582

Finding the Optimal Cardinality Value for Information Bottleneck Method
Gang Li, Dong Liu, Yiqing Tu, Yangdong Ye 594

A New Algorithm for Enumerating All Maximal Cliques in Complex Network
Li Wan, Bin Wu, Nan Du, Qi Ye, Ping Chen 606

Modeling and Mining the Rule Evolution
Ding Pan 618

Knowledge Reduction in Inconsistent Decision Tables
Qihe Liu, Leiting Chen, Jianzhong Zhang, Fan Min 626

Text Mining

Semantic Scoring Based on Small-World Phenomenon for Feature Selection in Text Mining
Chong Huang, Yonghong Tian, Tiejun Huang, Wen Gao 636

A Comparative Study on Text Clustering Methods
Yan Zheng, Xiaochun Cheng, Ronghuai Huang, Yi Man 644

Concept Based Text Classification Using Labeled and Unlabeled Data
Ping Gu, Qingsheng Zhu, Xiping He 652

Learning Semantic User Profiles from Text
M. Degemmis, P. Lops, G. Semeraro 661

Multimedia Mining

Audiovisual Integration for Racquet Sports Video Retrieval
Yaqin Zhao, Xianzhong Zhou, Guizhong Tang 673

A Correlation Approach for Automatic Image Annotation
David R. Hardoon, Craig Saunders, Sandor Szedmak, John Shawe-Taylor 681

Sequential Data Mining and Time Series Mining

Fast Discovery of Time-Constrained Sequential Patterns Using Time-Indexes
Ming-Yen Lin, Sue-Chen Hsueh, Chia-Wen Chang 693

Multi-dimensional Sequential Pattern Mining Based on Concept Lattice
Yang Jin, Wanli Zuo 702

Mining Time-Delayed Coherent Patterns in Time Series Gene Expression Data
Linjun Yin, Guoren Wang, Keming Mao, Yuhai Zhao 711

Mining Delay in Streaming Time Series of Industrial Process
Haijie Gu, Gang Rong 723

Segmental Semi-Markov Model Based Online Series Pattern Detection Under Arbitrary Time Scaling
Guangjie Ling, Yuntao Qian, Sen Jia 731

Diagnosis of Inverter Faults in PMSM DTC Drive Using Time-Series Data Mining Technique
Dan Sun, Jun Meng, Zongyuan He 741

Applications of Data Mining Time Series to Power Systems Disturbance Analysis
Jun Meng, Dan Sun, Zhiyong Li 749

Mining Compressed Sequential Patterns
Lei Chang, Dongqing Yang, Shiwei Tang, Tengjiao Wang 761

Effective Feature Preprocessing for Time Series Forecasting
Jun Hua Zhao, ZhaoYang Dong, Zhao Xu 769

On Similarity of Financial Data Series Based on Fractal Dimension
Jian-rong Hou, Hui Zhao, Pei Huang 782

Web Mining

A Hierarchical Model of Web Graph
Jie Han, Yong Yu, Chenxi Lin, Dingyi Han, Gui-Rong Xue 790

Web Scale Competitor Discovery Using Mutual Information
Rui Li, Shenghua Bao, Jin Wang, Yuanjie Liu, Yong Yu 798

Biomedical Mining

Co-expression Gene Discovery from Microarray for Integrative Systems Biology
Yutao Ma, Yonghong Peng 809

Cardiovascular Disease Diagnosis Method by Emerging Patterns
Heon Gyu Lee, Kiyong Noh, Bum Ju Lee, Ho-Sun Shon, Keun Ho Ryu 819

DNA Microarray Data Clustering by Hidden Markov Models and Bayesian Information Criterion
Phasit Charoenkwan, Aompilai Manorat, Jeerayut Chaijaruwanich, Sukon Prasitwattanaseree, Sakarindr Bhumiratana 827

Application of Factor Analysis on *Mycobacterium Tuberculosis* Transcriptional Responses for Drug Clustering, Drug Target, and Pathway Detections
Jeerayut Chaijaruwanich, Jamlong Khamphachua, Sukon Prasitwattanaseree, Saradee Warit, Prasit Palittapongarnpim 835

First Steps to an Audio Ontology-Based Classifier for Telemedicine
Cong Phuong Nguyen, Ngoc Yen Pham, Eric Castelli 845

Obstacles and Misunderstandings Facing Medical Data Mining
Ashkan Sami 856

SVM-Based Tumor Classification with Gene Expression Data
Shulin Wang, Ji Wang, Huowang Chen, Boyun Zhang 864

GEPCLASS: A Classification Rule Discovery Tool Using Gene Expression Programming
Wagner R. Weinert, Heitor S. Lopes 871

Advanced Applications

CBR-Based Knowledge Discovery on Results of Evolutionary Design of Logic Circuits
Shuguang Zhao, Mingying Zhao, Jin Li, Change Wang 881

Data Summarization Approach to Relational Domain Learning Based on Frequent Pattern to Support the Development of Decision Making
Rayner Alfred, Dimitar Kazakov 889

Extreme Value Dependence in Problems with a Changing Causation Structure
Marlon Núñez, Rafael Morales 899

A Study on Object Recognition Technology Using PCA in the Variable Illumination
Jong-Min Kim, Hwan-Seok Yang 911

Pattern Recurring in Three-Dimensional Graph Based on Data Mining
Yanbing Liu, Menghao Wang 919

Mining the Useful Skyline Set Based on the Acceptable Difference
Zhenhua Huang, Wei Wang 927

Modeling Information-Sharing Behaviors in BitTorrent System Based on Real Measurement
Jinkang Jia, Changjia Chen 934

Financial Distress Prediction Based on Similarity Weighted Voting CBR
Jie Sun, Xiao-Feng Hui 947

Customer Churn Prediction by Hybrid Model
Jae Sik Lee, Jin Chun Lee 959

Base Vector Selection for Kernel Matching Pursuit
Qing Li, Licheng Jiao 967

WaveSim Transform for Multi-channel Signal Data Mining Through Linear Regression PCA
R. Pradeep Kumar, P. Nagabhushan 977

Research on Query-by-Committee Method of Active Learning and Application
Yue Zhao, Ciwen Xu, Yongcun Cao 985

Traffic Management Genetic Algorithm Supporting Data Mining and QoS in Sensor Networks
Yantao Pan, Wei Peng, Xicheng Lu 992

Comparison of Data Pre-processing in Pattern Recognition of Milk Powder Vis/NIR Spectra
Haiyan Cen, Yidan Bao, Min Huang, Yong He 1000

Semi-automatic Hot Event Detection
Tingting He, Guozhong Qu, Siwei Li, Xinhui Tu, Yong Zhang, Han Ren 1008

Security and Privacy Issues

Profile-Based Security Against Malicious Mobile Agents
Hua Li, Glena Greene, Rafael Alonso 1017

A Comprehensive Categorization of DDoS Attack and DDoS Defense Techniques
Usman Tariq, ManPyo Hong, Kyung-suk Lhee 1025

Structural Analysis and Mathematical Methods for Destabilizing Terrorist Networks Using Investigative Data Mining
Nasrullah Memon, Henrik Legind Larsen 1037

Alert Correlation Analysis in Intrusion Detection
Moon Sun Shin, Kyeong Ja Jeong 1049

Spatial Data Mining

OSDM: Optimized Shape Distribution Method
Ashkan Sami, Ryoichi Nagatomi, Makoto Takahashi, Takeshi Tokuyama 1057

View-Angle of Spatial Data Mining
Shuliang Wang, Haning Yuan 1065

Streaming Data Mining

Maintaining Moving Sums over Data Streams
Tzu-Chiang Wu, Arbee L.P. Chen 1077

MFIS—Mining Frequent Itemsets on Data Streams
Zhi-jun Xie, Hong Chen, Cuiping Li 1085

Improving the Performance of Data Stream Classifiers by Mining Recurring Contexts
Yong Wang, Zhanhuai Li, Yang Zhang, Longbo Zhang, Yun Jiang 1094

Author Index 1107