

Table of Contents

Industrial Papers (Invited)

Network Data Mining and Analysis: The <i>NEMESIS</i> Project	1
<i>Minos Garofalakis, Rajeev Rastogi</i>	

Privacy Preserving Data Mining: Challenges and Opportunities	13
<i>Ramakrishnan Srikant</i>	

Survey Papers (Invited)

A Case for Analytical Customer Relationship Management	14
<i>Jaideep Srivastava, Jau-Hwang Wang, Ee-Peng Lim, San-Yih Hwang</i>	

On Data Clustering Analysis: Scalability, Constraints, and Validation	28
<i>Osmar R. Zaiane, Andrew Foss, Chi-Hoon Lee, Weinan Wang</i>	

Association Rules (I)

Discovering Numeric Association Rules via Evolutionary Algorithm	40
<i>Jacinto Mata, José-Luis Alvarez, José-Cristóbal Riquelme</i>	

Efficient Rule Retrieval and Postponed Restrict Operations for Association Rule Mining	52
<i>Jochen Hipp, Christoph Mangold, Ulrich Güntzer, Gholamreza Nakhaeizadeh</i>	

Association Rule Mining on Remotely Sensed Images Using P-trees	66
<i>Qin Ding, Qiang Ding, William Perrizo</i>	

On the Efficiency of Association-Rule Mining Algorithms	80
<i>Vikram Pudi, Jayant R. Haritsa</i>	

Classification (I)

A Function-Based Classifier Learning Scheme Using Genetic Programming	92
<i>Jung-Yi Lin, Been-Chian Chien, Tzung-Pei Hong</i>	

SNNB: A Selective Neighborhood Based Naïve Bayes for Lazy Learning . . .	104
<i>Zhipeng Xie, Wynne Hsu, Zongtian Liu, Mong Li Lee</i>	

A Method to Boost Naïve Bayesian Classifiers	115
<i>Lili Diao, Keyun Hu, Yuchang Lu, Chunyi Shi</i>	

Toward Bayesian Classifiers with Accurate Probabilities 123
Charles X. Ling, Huaqie Zhang

Interestingness

Pruning Redundant Association Rules Using Maximum Entropy
 Principle 135
Szymon Jaroszewicz, Dan A. Simovici

A Confidence-Lift Support Specification for Interesting Associations
 Mining 148
Wen-Yang Lin, Ming-Cheng Tseng, Ja-Hwung Su

Concise Representation of Frequent Patterns Based on Generalized
 Disjunction-Free Generators 159
Marzena Kryszkiewicz, Marcin Gajek

Mining Interesting Association Rules: A Data Mining Language 172
Show-Jane Yen, Yue-Shi Lee

The Lorenz Dominance Order as a Measure of Interestingness in KDD 177
Robert J. Hilderman

Sequence Mining

Efficient Algorithms for Incremental Update of Frequent Sequences 186
Minghua Zhang, Ben Kao, David Cheung, Chi-Lap Yip

DELISP: Efficient Discovery of Generalized Sequential Patterns by
 Delimited Pattern-Growth Technology 198
Ming-Yen Lin, Suh-Yin Lee, Sheng-Shun Wang

Self-Similarity for Data Mining and Predictive Modeling - A Case Study
 for Network Data 210
Jafar Adibi, Wei-Min Shen, Eaman Noorbakhsh

A New Mechanism of Mining Network Behavior 218
Shun-Chieh Lin, Shian-Shyong Tseng, Yao-Tsung Lin

Clustering

M-FastMap: A Modified FastMap Algorithm for Visual Cluster Validation
 in Data Mining 224
Michael Ng, Joshua Huang

An Incremental Hierarchical Data Clustering Algorithm Based on Gravity
 Theory 237
Chien-Yu Chen, Shien-Ching Hwang, Yen-Jen Oyang

Adding Personality to Information Clustering	251
<i>Ah-Hwee Tan, Hong Pan</i>	

Clustering Large Categorical Data	257
<i>François-Xavier Jollois, Mohamed Nadif</i>	

Web Mining

WebFrame: In Pursuit of Computationally and Cognitively Efficient Web Mining	264
<i>Tong Zheng, Yonghe Niu, Randy Goebel</i>	

Naviz: Website Navigational Behavior Visualizer	276
<i>Bowo Prasetyo, Iko Pramudiono, Katsumi Takahashi, Masaru Kitsuregawa</i>	

Optimal Algorithms for Finding User Access Sessions from Very Large Web Logs	290
<i>Zhixiang Chen, Ada Wai-Chee Fu, Frank Chi-Hung Tong</i>	

Automatic Information Extraction for Multiple Singular Web Pages	297
<i>Chia-Hui Chang, Shih-Chien Kuo, Kuo-Yu Hwang, Tsung-Hsin Ho, Chih-Lung Lin</i>	

Association Rules (II)

An Improved Approach for the Discovery of Causal Models via MML	304
<i>Honghua Dai, Gang Li</i>	

SETM*-MaxK: An Efficient SET-Based Approach to Find the Largest Itemset	316
<i>Ye-In Chang, Yu-Ming Hsieh</i>	

Discovery of Ordinal Association Rules	322
<i>Sylvie Guillaume</i>	

Value Added Association Rules	328
<i>T.Y. Lin, Y.Y. Yao, E. Louie</i>	

Top Down FP-Growth for Association Rule Mining	334
<i>Ke Wang, Liu Tang, Jiawei Han, Junqiang Liu</i>	

Semi-structure & Concept Mining

Discovery of Frequent Tag Tree Patterns in Semistructured Web Documents	341
<i>Tetsuhiro Miyahara, Yusuke Suzuki, Takayoshi Shoudai, Tomoyuki Uchida, Kenichi Takahashi, Hiroaki Ueda</i>	

Extracting Characteristic Structures among Words in Semistructured Documents 356
Kazuyoshi Furukawa, Tomoyuki Uchida, Kazuya Yamada, Tetsuhiro Miyahara, Takayoshi Shoudai, Yasuaki Nakamura

An Efficient Algorithm for Incremental Update of Concept Space 368
Felix Cheung, Ben Kao, David Cheung, Chi-Yuen Ng

Data Warehouse and Data Cube

Efficient Constraint-Based Exploratory Mining on Large Data Cubes 381
Cuiping Li, Shengen Li, Shan Wang, Xiaoyong Du

Efficient Utilization of Materialized Views in a Data Warehouse 393
Don-Lin Yang, Man-Lin Huang, Ming-Chuan Hung

Bio-Data Mining

Mining Interesting Rules in Meningitis Data by Cooperatively Using GDT-RS and RSBR 405
Ning Zhong, Juzhen Dong

Evaluation of Techniques for Classifying Biological Sequences 417
Mukund Deshpande, George Karypis

Efficiently Mining Gene Expression Data via Integrated Clustering and Validation Techniques 432
Vincent S.M. Tseng, Ching-Pin Kao

Classification (II)

Adaptive Generalized Estimation Equation with Bayes Classifier for the Job Assignment Problem 438
Yulan Liang, King-Ip Lin, Arpad Kelemen

GEC: An Evolutionary Approach for Evolving Classifiers 450
William W. Hsu, Ching-Chi Hsu

An Efficient Single-Scan Algorithm for Mining Essential Jumping Emerging Patterns for Classification 456
Hongjian Fan, Ramamohanarao Kotagiri

A Method to Boost Support Vector Machines 463
Lili Diao, Keyun Hu, Yuchang Lu, Chunyi Shi

Temporal Mining

Distribution Discovery: Local Analysis of Temporal Rules 469
Xiaoming Jin, Yuchang Lu, Chunyi Shi

News Sensitive Stock Trend Prediction	481
<i>Gabriel Pui Cheong Fung, Jeffrey Xu Yu, Wai Lam</i>	
User Profiling for Intrusion Detection Using Dynamic and Static Behavioral Models	494
<i>Dit-Yan Yeung, Yuxin Ding</i>	
Classification (III)	
Incremental Extraction of Keyterms for Classifying Multilingual Documents in the Web	506
<i>Lee-Feng Chien, Chien-Kang Huang, Hsin-Chen Chiao, Shih-Jui Lin</i>	
k -nearest Neighbor Classification on Spatial Data Streams Using P-trees ..	517
<i>Maleq Khan, Qin Ding, William Perrizo</i>	
Interactive Construction of Classification Rules	529
<i>Jianchao Han, Nick Cercone</i>	
Outliers, Missing Data, and Causation	
Enhancing Effectiveness of Outlier Detections for Low Density Patterns ...	535
<i>Jian Tang, Zhixiang Chen, Ada Wai-chee Fu, David W. Cheung</i>	
Cluster-Based Algorithms for Dealing with Missing Values	549
<i>Yoshikazu Fujikawa, TuBao Ho</i>	
Extracting Causation Knowledge from Natural Language Texts	555
<i>Ki Chan, Boon-Toh Low, Wai Lam, Kai-Pui Lam</i>	
Mining Relationship Graphs for Effective Business Objectives	561
<i>Kok-Leong Ong, Wee-Keong Ng, Ee-Peng Lim</i>	
Author Index	567



<http://www.springer.com/978-3-540-43704-8>

Advances in Knowledge Discovery and Data Mining
6th Pacific-Asia Conference, PAKDD 2002, Taipei,
Taiwan, May 6-8, 2002. Proceedings
Cheng, M.-S.; Yu, P.S.; Liu, B. (Eds.)
2002, XIV, 570 p., Softcover
ISBN: 978-3-540-43704-8