

2011 IEEE 11th International Conference on Data Mining Workshops

(ICDMW 2011)

**Vancouver, British Columbia, Canada
11 December 2011**

Pages 1-627



IEEE Catalog Number: CFP1156B-PRT
ISBN: 978-1-4673-0005-6

2011 11th IEEE International Conference on Data Mining Workshops

ICDMW 2011

Table of Contents

Message from the Conference Chairs	xviii
Message from the Workshop Co-chairs	xix
Preface to the Third IEEE ICDM Workshop on Knowledge Discovery from Climate Data	xxi
Preface to the Sentiment Elicitation from Natural Text for Information Retrieval and Extraction Workshop	xxiii
Preface to the International Workshop on Knowledge Discovery Using Cloud and Distributed Computing Platforms	xxv
Preface to the Sixth Workshop on Optimization Based Techniques for Emerging Data Mining Problems	xxvii
Preface to the Data Mining Technologies for Computational Collective Intelligence Workshop	xxix
Preface to the Workshop on Mining Communities and People Recommenders	xxxi
Preface to the Workshop on Domain Driven Data Mining	xxxiii
Preface to the Fifth International Workshop on Mining Multiple Information Sources	xxxv
Preface to the Handling Concept Drift and Reoccurring Contexts in Adaptive Information Systems Workshop	xxxvi
Preface to the Third International Workshop on Privacy Aspects of Data Mining	xxxviii
Preface to the International Workshop on Learning and Data Mining for Robots	xl
Preface to the Fourth Workshop on Data Mining Case Studies	xli
Preface to the Large Scale Visual Analytics Workshop	xliii
Preface to the International Workshop on Spatial and Spatio-Temporal Data Mining	xliv
Preface to the International Workshop on Data Mining in Networks	xlvi
Preface to the Second Workshop on Biological Data Mining and Its Applications in Healthcare	l
Preface to the Declarative Pattern Mining Workshop	liii

Preface to the Contrast Data Mining and Applications Workshop	lv
Preface to the Data Mining for Service Workshop	lvi
Preface to the IEEE ICDM 2011 PhD Forum	lvii
Preface to the ICDM 2011 Demo Session	lvix
Sponsors	lx

2011 Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes, and Impacts

Sparse Group Lasso for Regression on Land Climate Variables	1
<i>Soumyadeep Chatterjee, Arindam Banerjee, Snigdhanshu Chatterjee, and Auroop R. Ganguly</i>	
Community Dynamics and Analysis of Decadal Trends in Climate Data	9
<i>William Hendrix, Isaac K. Tetteh, Ankit Agrawal, Fredrick Semazzi, Wei-keng Liao, and Alok Choudhary</i>	
Modeling Unreliable Data and Sensors: Using F-measure Attribute Performance with Test Samples from Low-Cost Sensors	15
<i>Vasanth Iyer and S. Sitharama Iyengar</i>	
Evaluating Hurricane Intensity Prediction Techniques in Real Time	23
<i>Vladimir Jovanovic, Margaret H. Dunham, Michael Hahsler, and Yu Su</i>	
Spatio-Temporal Mining of Core Regions: Study of Rainfall Patterns in Monsoonal India	30
<i>Kollukuduru Sravanthi and K.S. Rajan</i>	
Parallel Kriging Analysis for Large Spatial Datasets	38
<i>Wei Zhuo, Prabhat, Chris Paciorek, Cari Kaufman, and Wes Bethel</i>	

ICDM SENTIRE: Sentiment Elicitation from Natural Text for Information Retrieval and Extraction

Multilingual Sentiment Analysis Using Latent Semantic Indexing and Machine Learning	45
<i>Brett W. Bader, W. Philip Kegelmeyer, and Peter A. Chew</i>	
Deriving Insights from National Happiness Indices	53
<i>Anthony Brew, Derek Greene, Daniel Archambault, and Pádraig Cunningham</i>	
Improving Sentiment Classification Using Feature Highlighting and Feature Bagging	61
<i>Lin Dai, Hechun Chen, and Xuemei Li</i>	
STARLET: Multi-document Summarization of Service and Product Reviews with Balanced Rating Distributions	67
<i>Giuseppe Di Fabbrizio, Ahmet Aker, and Robert Gaizauskas</i>	
Machine Reading for Notion-Based Sentiment Mining	75
<i>Roula Hobeica, Hazem Hajj, and Wassim El Hajj</i>	
Multi-aspect Sentiment Analysis with Topic Models	81
<i>Bin Lu, Myle Ott, Claire Cardie, and Benjamin K. Tsou</i>	
AQA: Aspect-based Opinion Question Answering	89
<i>Samaneh Moghaddam and Martin Ester</i>	

Discourse Structure and Sentiment	97
<i>Livia Polanyi and Martin van den Berg</i>	
Longitudinal Sales Responses with Online Reviews	103
<i>Xiaojun Quan, Liu Wenyin, and Wenyu Dou</i>	
Fine-Grained Opinion Mining Using Conditional Random Fields	109
<i>Shabnam Shariaty and Samaneh Moghaddam</i>	
Detecting General Opinions from Customer Surveys	115
<i>Evgeny A. Stepanov and Giuseppe Riccardi</i>	
Mining Opinion Attributes from Texts Using Multiple Kernel Learning	123
<i>Aleksander Wawer</i>	
SES: Sentiment Elicitation System for Social Media Data	129
<i>Kunpeng Zhang, Yu Cheng, Yusheng Xie, Daniel Honbo, Ankit Agrawal, Diana Palsetia, Kathy Lee, Wei-keng Liao, and Alok Choudhary</i>	

KDCloud 2011: International Workshop on Knowledge Discovery Using Cloud and Distributed Computing Platforms

Detecting Abnormal Machine Characteristics in Cloud Infrastructures	137
<i>Kanishka Bhaduri, Kamalika Das, and Bryan L. Matthews</i>	
Frequent Pairs in Data Streams: Exploiting Parallelism and Skew	145
<i>Andrea Campagna, Konstantin Kutzkov, and Rasmus Pagh</i>	
Epidemic K-Means Clustering	151
<i>Giuseppe Di Fatta, Francesco Blasa, Simone Cafiero, and Giancarlo Fortino</i>	
A Large Scale URL Verification Pipeline Using Hadoop	159
<i>Songtao Guo and Jianxiong Dong</i>	
Latency Minimizing Tasking for Information Processing Systems	167
<i>James Horey and Brent Lagesse</i>	
LI-MR: A Local Iteration Map/Reduce Model and Its Application to Mine Community Structure in Large-Scale Networks	174
<i>QiuHong Li, Zhihui Wang, Wei Wang, Yimin Liu, Peng Wang, and Tao Yu</i>	
Finding the ‘Needle’: Locating Interesting Nodes Using the K-shortest Paths Algorithm in MapReduce	180
<i>Christopher McCubbin, Bryan Perozzi, Andrew Levine, and Abdul Rahman</i>	
The Deployment of MML for Data Analytics over the Cloud	188
<i>Jonathan Tancer and Aparna S. Varde</i>	
Temporal Distributed Learning with Heterogeneous Data Using Gaussian Mixtures	196
<i>Dean Teffer, Amanda Hutton, and Joydeep Ghosh</i>	

OEDM 2011: The 6th Workshop on Optimization Based Techniques for Emerging Data Mining Problems

Post Mining of Multiple Criteria Linear Programming Classification Model for Actionable Knowledge in Credit Card Churning Management	204
<i>Yibing Chen, Lingling Zhang, and Yong Shi</i>	
Learning to Group Web Text Incorporating Prior Information	212
<i>Yu Cheng, Kunpeng Zhang, Yusheng Xie, Ankit Agrawal, Wei-keng Liao, and Alok Choudhary</i>	
Meta-learning for Selecting a Multi-label Classification Algorithm	220
<i>Lena Chekina, Lior Rokach, and Bracha Shapira</i>	
Efficient Iterative Semi-supervised Classification on Manifold	228
<i>Mehrdad Farajtabar, Hamid R. Rabiee, Amirreza Shaban, and Ali Soltani-Farani</i>	
Boosting Unsupervised Additive Clustering Using Cluster-Wise Optimization and Multi-label Learning	236
<i>Stephen L. France and Ahmed Abbasi</i>	
Active Learning from Positive and Unlabeled Data	244
<i>Alireza Ghasemi, Hamid R. Rabiee, Mohsen Fadaee, Mohammad T. Manzuri, and Mohammad H. Rohban</i>	
Twitter Trending Topic Classification	251
<i>Kathy Lee, Diana Palsetia, Ramanathan Narayanan, Md. Mostofa Ali Patwary, Ankit Agrawal, and Alok Choudhary</i>	
Temporal Cross-Selling Optimization Using Action Proxy-Driven Reinforcement Learning	259
<i>Nan Li and Naoki Abe</i>	
Entity Resolution with Attribute and Connection Graph	267
<i>Lingfeng Niu, Jianmin Wu, and Yong Shi</i>	
Integrating Pairwise Constraints into Clustering Algorithms: Optimization-Based Approaches	272
<i>Jacques-Henri Sublemontier, Lionel Martin, Guillaume Cleuziou, and Matthieu Exbrayat</i>	
Classification for Orange Varieties Using Near Infrared Spectroscopy	280
<i>Warawut Suphamitmongkol, Guangli Nie, Rong Liu, and Yong Shi</i>	
Parametric Characterization of Multimodal Distributions with Non-gaussian Modes	286
<i>Ashutosh Tewari, Michael J. Giering, and Arvind Raghunathan</i>	
Kernel-Based Clustering with Automatic Cluster Number Selection	293
<i>Chang-Dong Wang, Jian-Huang Lai, and Dong Huang</i>	
A Novel Co-clustering Method with Intra-similarities	300
<i>Jian-Sheng Wu, Jian-Huang Lai, and Chang-Dong Wang</i>	
Stream Prediction Using Representative Episode Rules	307
<i>Huasheng Zhu, Peng Wang, Wei Wang, and Baile Shi</i>	

DMCCI 2011: Data Mining Technologies for Computational Collective Intelligence

Isanette: A Common and Common Sense Knowledge Base for Opinion Mining	315
<i>Erik Cambria, Yangqiu Song, Haixun Wang, and Amir Hussain</i>	
Mobile Phone Graph Evolution: Findings, Model and Interpretation	323
<i>Siyuan Liu, Lei Li, Christos Faloutsos, and Lionel M. Ni</i>	
Exploring Support Vector Machines and Random Forests to Detect Advanced Fee Fraud Activities on Internet	331
<i>Abiodun Modupe, Oludayo O. Olugbara, and Sunday O. Ojo</i>	
Retweet Modeling Using Conditional Random Fields	336
<i>Huan-Kai Peng, Jiang Zhu, Dongzhen Piao, Rong Yan, and Ying Zhang</i>	
SLPA: Uncovering Overlapping Communities in Social Networks via a Speaker-Listener Interaction Dynamic Process	344
<i>Jierui Xie, Boleslaw K. Szymanski, and Xiaoming Liu</i>	

COMMPER 2011: Workshop on Mining Communities and People Recommenders

Employing Team Composition Strategies for Recommending Teams	350
<i>Michele Brocco and Yonata Andrelo Asikin</i>	
Random Walk Based Resource Allocation: Predicting and Recommending Links in Cross-Operator Mobile Communication Networks	358
<i>Yuxiao Dong, Qing Ke, Jun Rao, Bai Wang, and Bin Wu</i>	
An Effective Expert Team Formation in Social Networks Based on Skill Grading	366
<i>Farnoush Farhadi, Maryam Sorkhi, Sattar Hashemi, and Ali Hamzeh</i>	
Time-Aware Ranking in Dynamic Citation Networks	373
<i>Rumi Ghosh, Tsung-Ting Kuo, Chun-Nan Hsu, Shou-De Lin, and Kristina Lerman</i>	
A Diffusion of Innovation-Based Closeness Measure for Network Associations	381
<i>Reihaneh Rabbany Khorasgani and Osmar R. Zaïane</i>	

DDDM 2011: Workshop on Domain Driven Data Mining

Improving Energy Use Forecast for Campus Micro-grids Using Indirect Indicators	389
<i>Saima Aman, Yogesh Simmhan, and Viktor K. Prasanna</i>	
FAARM: Frequent Association Action Rules Mining Using FP-Tree	398
<i>Djellel Eddine Difallah, Ryan G. Benton, Vijay Raghavan, and Tom Johnsten</i>	
Certainty-Enhanced Active Learning for Improving Imbalanced Data Classification	405
<i>Jui Hsi Fu and Sing Ling Lee</i>	
Automatic Cleaning and Linking of Historical Census Data Using Household Information	413
<i>Zhichun Fu, Peter Christen, and Mac Boot</i>	
Mining Infrequent Causal Associations in Electronic Health Databases	421
<i>Yanqing Ji, Hao Ying, John Tran, Peter Dews, Ayman Mansour, and R. Michael Massanari</i>	
A Novel Knowledge-Based Twin Support Vector Machine	429
<i>Xu Chan Ju and Ying Jie Tian</i>	

Semi-supervised Failure Prediction for Oil Production Wells	434
<i>Yintao Liu, Ke-Thia Yao, Shuping Liu, Cauligi S. Raghavendra, Oluwafemi Balogun, and Lanre Olabinjo</i>	
Automatic Training Data Cleaning for Text Classification	442
<i>Hassan H. Malik and Vikas S. Bhardwaj</i>	
Revealing Cluster Formation over Huge Volatile Robotic Data	450
<i>Nikos Mitsou, Irene Ntoutsis, Dirk Wollherr, Costas Tzafestas, and Hans-Peter Kriegel</i>	
Domain-Driven Data Mining in Human Resource Management: A Review	458
<i>Franca Piazza and Stefan Strohmeier</i>	
CAPRE: A New Methodology for Product Recommendation Based on Customer Actionability and Profitability	466
<i>Thomas Piton, Julien Blanchard, and Fabrice Guillet</i>	
Domain-Specific Adaptation of a Partial Least Squares Regression Model for Loan Defaults Prediction	474
<i>Balaji Vasan Srinivasan, Nathan Gnanasambandam, Shi Zhao, and Raj Minhas</i>	
Agent Assignment for Process Management: Competency-Driven Dynamic Resource Management Methodology	480
<i>Ramzan Talib, Bernhard Volz, and Stefan Jablonski</i>	
 MMIS 2011: The 5th International Workshop on Mining Multiple Information Sources	
Relational Fuzzy Clustering with Multiple Kernels	488
<i>Naouel Baili and Hichem Frigui</i>	
Cross-Domain Recommender Systems	496
<i>Paolo Cremonesi, Antonio Tripodi, and Roberto Turrin</i>	
Approximate Record Matching Using Hash Grams	504
<i>Mohammed Gollapalli, Xue Li, Ian Wood, and Guido Governatori</i>	
Multitask Multiclass Support Vector Machines	512
<i>You Ji and Shiliang Sun</i>	
Indexing Faces in Broadcast News Video Archives	519
<i>Duy-Dinh Le and Shin'ichi Satoh</i>	
Greedy Regularized Least-Squares for Multi-task Learning	527
<i>Pekka Naula, Tapio Pahikkala, Antti Airola, and Tapio Salakoski</i>	
Using Modified Multivariate Bag-of-Words Models to Classify Physiological Data	534
<i>Patricia Ordóñez, Tom Armstrong, Tim Oates, and Jim Fackler</i>	
Transportability of Causal and Statistical Relations: A Formal Approach	540
<i>Judea Pearl and Elias Bareinboim</i>	
A Tag-Based Hybrid Music Recommendation System Using Semantic Relations and Multi-domain Information	548
<i>Ipek Tatli and Ayse Nur Birtürk</i>	

HaCDAIS 2011: Handling Concept Drift and Reoccurring Contexts in Adaptive Information Systems

Unifying Change—Towards a Framework for Detecting the Unexpected	555
<i>Iris Adä and Michael R. Berthold</i>	
Change Mining of Customer Profiles Based on Transactional Data	560
<i>Edward Apeh and Bogdan Gabrys</i>	
Detecting Mean Changes in Data Streams	568
<i>Murad Badarna and Ran Wolff</i>	
What’s Your Current Stress Level? Detection of Stress Patterns from GSR Sensor Data	573
<i>Jorn Bakker, Mykola Pechenizkiy, and Natalia Sidorova</i>	
Interpretable, Online Soft-Sensors for Process Control	581
<i>Mark Eastwood and Petr Kadlec</i>	
Pool and Accuracy Based Stream Classification: A New Ensemble Algorithm on Data Stream Classification Using Recurring Concepts Detection	588
<i>Mohammad Javad Hosseini, Zahra Ahmadi, and Hamid Beigy</i>	
Classification in Presence of Drift and Latency	596
<i>Georg Kreml and Vera Hofer</i>	
Drift Detection Using Uncertainty Distribution Divergence	604
<i>Patrick Lindstrom, Brian Mac Namee, and Sarah Jane Delany</i>	
Interpretability of Sudden Concept Drift in Medical Informatics Domain	609
<i>Gregor Stiglic and Peter Kokol</i>	

PADM 2011: Third International Workshop on Privacy Aspects of Data Mining

An Application of Differentially Private Linear Mixed Modeling	614
<i>John M. Abowd and Matthew J. Schneider</i>	
enList: Automatically Simplifying Privacy Policies	620
<i>Rajesh Bejugam and Kristen LeFevre</i>	
Privacy Preserving GWAS Data Sharing	628
<i>Stephen E. Fienberg, Aleksandra Slavković, and Caroline Uhler</i>	
Data Mining and Privacy of Personal Behaviour Types in Smart Grid	636
<i>George Kalogridis and Stojan Z. Denic</i>	
Fairness-aware Learning through Regularization Approach	643
<i>Toshihiro Kamishima, Shotaro Akaho, and Jun Sakuma</i>	
Applicability of Regression-Tree-Based Synthetic Data Methods for Business Data	651
<i>Joo-Ho Lee, In-Yong Kim, and Christine M. O’Keefe</i>	
Preventing Identity Disclosure in Hypergraphs	659
<i>Yidong Li and Hong Shen</i>	
Rating: Privacy Preservation for Multiple Attributes with Different Sensitivity Requirements	666
<i>Jinfei Liu, Jun Luo, and Joshua Zhexue Huang</i>	

Privacy Preserving Outlier Detection Using Locality Sensitive Hashing	674
<i>Nisarg Raval, Madhuchand Rushi Pillutla, Piyush Bansal, Kannan Srinathan, and C.V. Jawahar</i>	

LEMIR 2011: International Workshop on Learning and Data Mining for Robots

Implementing Camshift on a Mobile Robot for Person Tracking and Pursuit	682
<i>Somar Boubou, Asuki Kouno, and Einoshin Suzuki</i>	
Lifted-Rollout for Approximate Policy Iteration of Markov Decision Process	689
<i>Wang-Zhou Dai, Yang Yu, and Zhi-Hua Zhou</i>	
Adaptive Windowing for Online Learning from Multiple Inter-related Data Streams	697
<i>Elena Ikonomovska, Kurt Driessens, Sašo Džeroski, and João Gama</i>	
Characterizing Anomalous Behaviors and Revising Robotic Controllers	705
<i>David Meunier, Michele Sebag, and Shin Ando</i>	

DMCS 2011: The Fourth Workshop on Data Mining Case Studies

Integrating Collaborative Filtering and Search-Based Techniques for Personalized Online Product Recommendation	711
<i>Noraswaliza Abdullah, Yue Xu, and Shlomo Geva</i>	
“The 100 Most Influential Persons in History”: A Data Mining Perspective	719
<i>Noora Mohammad Al-Naimi and Khaled Bashir Shaban</i>	
Evolutionary Algorithms for Selecting the Architecture of a MLP Neural Network: A Credit Scoring Case	725
<i>Alejandro Correa Bahnsen and Andres M. Gonzalez</i>	
A/B Testing at SweetIM: The Importance of Proper Statistical Analysis	733
<i>Slava Borodovsky and Saharon Rosset</i>	
Detection of Anomalous Particles from the Deepwater Horizon Oil Spill Using the SIPPER3 Underwater Imaging Platform	741
<i>Sergiy Fefilatyeu, Kurt Kramer, Lawrence Hall, Dmitry Goldgof, Rangachar Kasturi, Andrew Remsen, and Kendra Daly</i>	
Dynamic Loan Service Monitoring Using Segmented Hidden Markov Models	749
<i>Haengju Lee, Nathan Gnanasambandam, Raj Minhas, and Shi Zhao</i>	
An Analysis of Treatment Patterns: A Case Study on Carpal Tunnel Syndrome	755
<i>Ching Lien and Silvia Figueira</i>	
Evaluation and Extension to the Duckworth Lewis Method: A Dual Application of Data Mining Techniques	763
<i>Viraj Phanse and Sourabh Deorah</i>	
An Augmented Vector Space Information Retrieval for Recovering Requirements Traceability	771
<i>Yoshihisa Udagawa</i>	
Crime Forecasting Using Data Mining Techniques	779
<i>Chung-Hsien Yu, Max W. Ward, Melissa Morabito, and Wei Ding</i>	

LSVA 2011: Large Scale Visual Analytics

Clustering Based Fast Low-Rank Approximation for Large-Scale Graph	787
<i>Wei Chen, Ming Shao, and Yun Fu</i>	
Efficient Indexing for Mobile Image Retrieval	793
<i>Deying Feng, Jie Yang, and Cheng Yang</i>	
Online Heterogeneous Feature Fusion for Visual Recognition	799
<i>Shuangping Huang, Lianwen Jin, and Xiaoxin Wei</i>	
FpMapViz: A Space-Filling Visualization for Frequent Patterns	804
<i>Carson Kai-Sang Leung, Fan Jiang, and Pourang P. Irani</i>	
Wisdom of Crowds: Single Image Super-resolution from the Web	812
<i>Jun Li and Dacheng Tao</i>	
Learning to Combine Ad-hoc Ranking Functions for Image Retrieval	817
<i>Yangxi Li, Bo Geng, Chao Zhou, and Chao Xu</i>	
From Videos to Places: Geolocating the World's Videos	823
<i>Jasper Snoek, Luciano Sbaiz, and Hrishikesh Aradhya</i>	
Visual Analysis of Bipartite Networks	833
<i>Supaporn Spanurattana and Tsuyoshi Murata</i>	
Incremental Support Vector Clustering	839
<i>Chang-Dong Wang, Jian-Huang Lai, and Dong Huang</i>	
Large-scale Outdoor Scene Classification by Boosting a Set of Highly Discriminative and Low Redundant Graphlets	847
<i>Luming Zhang, Mingli Song, Xiaoyu Deng, Jiajun Bu, and Chun Chen</i>	

SSTDM 2011: International Workshop on Spatial and Spatio-Temporal Data Mining

Incremental Maintenance of Topological Patterns in Spatial-Temporal Database	853
<i>Yi-Cheng Chen, Chao-Ying Wu, and Suh-Yin Lee</i>	
Efficiently Mining Dynamic Zonal Co-location Patterns Based on Maximal Co-locations	861
<i>Bi-Ru Dai and Meng-Yan Lin</i>	
High-Resolution Urban Image Classification Using Extended Features	869
<i>Ranga Raju Vatsavai</i>	
Trended DTW Based on Piecewise Linear Approximation for Time Series Mining	877
<i>Lei Sun, Yujia Yang, and Wenhua Liu</i>	
Data Mining Cancer Registries: Retrospective Surveillance of Small Area Time Trends in Cancer Incidence Using BaySTDetect	885
<i>Guangquan Li, Sylvia Richardson, Lea Fortunato, Ismail Ahmed, Anna Hansell, Mireille Toledano, and Nicky Best</i>	
Origin/Destination-estimation Using Cellular Network Data	891
<i>Erik Mellegård, Simon Moritz, and Mohamed Zahoor</i>	
Linearly-Combined Web Sensors for Spatio-temporal Data Extraction from the Web	897
<i>Shun Hattori</i>	

Spatial and Temporal Analysis of Planet Scale Vehicular Imagery Data905
Gautam S. Thakur, Pan Hui, Hamed Ketabdar, and Ahmed Helmy

New Object Detection for On-board Robot Vision by Lifting Complex Wavelet Transforms911
Shigeru Takano and Einoshin Suzuki

DaMNet 2011: International Workshop on Data Mining in Networks

Prophet—A Link-Predictor to Learn New Rules on NELL917
Ana Paula Appel and Estevam Rafael Hruschka Junior

Temporal Scale of Processes in Dynamic Networks925
Rajmonda Sulo Caceres, Tanya Berger-Wolf, and Robert Grossman

Working for Influence: Effect of Network Density and Modularity on Diffusion in Networks933
Habiba Habiba and Tanya Berger-Wolf

Automatically Spotting Information-Rich Nodes in Graphs941
Xiao He, Jing Feng, and Claudia Plant

Stochastic Network Motif Detection in Social Media949
Kai Liu, William K. Cheung, and Jiming Liu

Imputation of Missing Links and Attributes in Longitudinal Social Surveys957
Vladimir Ouzienko and Zoran Obradovic

Poll: A Citation Text Based System for Identifying High-Impact Contributions of an Article965
Lalith Polepeddi, Ankit Agrawal, and Alok Choudhary

Diffusion in Networks with Overlapping Community Structure969
Fergal Reid and Neil Hurley

Scalable Link Prediction on Multidimensional Networks979
Giulio Rossetti, Michele Berlingerio, and Fosca Giannotti

Collaborative Filtering for Coordinated Monitoring in Sensor Networks987
Janne Toivola and Jaakko Hollmén

BioDM 2011: The Second Workshop on Biological Data Mining and Its Applications in Healthcare

Identifying HotSpots in Lung Cancer Data Using Association Rule Mining995
Ankit Agrawal and Alok Choudhary

Confident Surgical Decision Making in Temporal Lobe Epilepsy by Heterogeneous Classifier
Ensembles1003
*Shobeir Fakhraei, Hamid Soltanian-Zadeh, Kourosh Jafari-Khouzani, Kost Elisevich,
and Farshad Fotouhi*

From Sequences to Papers: An Information Retrieval Exercise1010
Célia Talma Gonçalves, Rui Camacho, and Eugénio Oliveira

Transduction of Semi-supervised Regression Targets in Survival Analysis for Medical Prognosis1018
Faisal M. Khan and Qiuhua Liu

Analyzing Trends of Hospital Length of Stay Using Phase-Type Distributions1026
Truc Viet Le, Chee Keong Kwoh, Eng Soon Teo, and Kheng Hock Lee

Graph Based Classification of MRI Data Based on the Ventricular System	1034
<i>S. Seth Long and Lawrence B. Holder</i>	
Medical Data Mining for Early Deterioration Warning in General Hospital Wards	1042
<i>Yi Mao, Yixin Chen, Gregory Hackmann, Minmin Chen, Chenyang Lu, Marin Kollef, and Thomas C. Bailey</i>	
Mining Protein Sequence Databases for Remote Homologues That Can Display Considerable Domain Length Variations	1050
<i>Eshita Mutt, Abhijit Mitra, and R. Sowdhamini</i>	
Brain Tumor Pathological Area Delimitation through Non-negative Matrix Factorization	1058
<i>Sandra Ortega-Martorell, Paulo J.G. Lisboa, Alfredo Vellido, Rui V. Simões, Margarida Julià-Sapé, and Carles Arús</i>	
<i>In Vivo</i> and <i>In Silico</i> Evidence: Hippocampal Cholesterol Metabolism Decreases with Aging and Increases with Alzheimers Disease—Modeling Brain Aging and Disease	1064
<i>Clyde F. Phelix, Richard G. LeBaron, Dawnlee J. Roberson, Rosa E. Villanueva, Greg Villareal, Omid B. Rahimi, Sandra Siedlak, Xiongwei Zhu, and George Perry</i>	
A Three-Step Validation Following Genome-Wide Data Mining for Myosin Family Members Improves Search Efficiency	1071
<i>Divya P. Syamaladevi, S. Kalaimathy, Naseer Pasha, S. Subramonian, and R. Sowdhamini</i>	
FTCluster: Efficient Mining Fault-Tolerant Biclusters in Microarray Dataset	1075
<i>Miao Wang, Xuequn Shang, Miao Miao, Zhanhuai Li, and Wenbin Liu</i>	
 DPM 2011: Declarative Pattern Mining	
Mining of <i>EL</i> -GCIs	1083
<i>Daniel Borchmann and Felix Distel</i>	
A Constraint Programming Approach for Enumerating Motifs in a Sequence	1091
<i>Emmanuel Coquery, Saïd Jabbour, and Lakhdar Saïs</i>	
CLIM: Closed Inclusion Dependency Mining in Databases	1098
<i>Fabien De Marchi</i>	
Declarative Heuristic Search for Pattern Set Mining	1104
<i>Tias Guns, Siegfried Nijssen, Albrecht Zimmermann, and Luc De Raedt</i>	
A Constraint-Based Language for Declarative Pattern Discovery	1112
<i>Jean-Philippe Métivier, Patrice Boizumault, Bruno Crémilleux, Mehdi Khiari, and Samir Loudni</i>	
Constraint-Based Pattern Mining in Multi-relational Databases	1120
<i>Siegfried Nijssen, Aïda Jiménez, and Tias Guns</i>	
 ContrastDM 2011: Contrast Data Mining and Applications	
Visually Contrast Two Collections of Frequent Patterns	1128
<i>Christopher L. Carmichael, Yaroslav Hayduk, and Carson Kai-Sang Leung</i>	
A New Algorithm Based on Shared Pattern-Tree to Mine Shared Emerging Patterns	1136
<i>Xiangtao Chen and Lijuan Lu</i>	

Overview of Contrast Data Mining as a Field and Preview of an Upcoming Book	1141
<i>Guozhu Dong and James Bailey</i>	
Correlation and Contrast Link Formation Patterns in a Time Evolving Graph	1147
<i>Tomonobu Ozaki and Minoru Etoh</i>	
Summarizing Contrasts by Recursive Pattern Mining	1155
<i>Arnaud Soulet, Bruno Crémilleux, and Marc Plantevit</i>	
DMS 2011: Data Mining for Service	
RFM Variables Revisited Using Quantile Regression	1163
<i>Michel Ballings, Dries Benoit, and Dirk Van den Poel</i>	
Analysis of Residence Time in Shopping Using RFID Data—An Application of the Kernel Density Estimation to RFID	1170
<i>Shinya Miyazaki, Takashi Washio, and Katsutoshi Yada</i>	
An Information Extraction Method from Different Structural Web Sites by Word Distances between a User Instantiated Label and Similar Entity	1177
<i>Daisuke Nakajima, Yuki Mitsui, Masaki Samejima, and Masanori Akiyoshi</i>	
Visualization of Hospital Services Using Data Mining Methods	1183
<i>Shusaku Tsumoto, Shoji Hirano, and Yuko Tsumoto</i>	
Estimating Post-Event Seller Productivity Profiles in Dynamic Sales Organizations	1191
<i>Kush R. Varshney, Moninder Singh, Mayank Sharma, and Aleksandra Mojsilovic</i>	
PhD Forum	
On Efficient Distance-Based Similarity Search	1199
<i>Jianquan Liu, Hanxiong Chen, Kazutaka Furuse, Hiroyuki Kitagawa, and Jeffrey Xu Yu</i>	
Active Learning of Transfer Relationships for Multiple Related Bayesian Network Structures	1203
<i>Diane Oyen</i>	
Understanding and Exploiting the Connections between NMF and SVM	1207
<i>Vamsi K. Potluru</i>	
Community Evolution in Dynamic Social Networks—Challenges and Problems	1211
<i>Mansoureh Takaffoli</i>	
Nearest Neighbor Voting in High-Dimensional Data: Learning from Past Occurrences	1215
<i>Nenad Tomašev and Dunja Mladenic</i>	
Towards a Framework for Detecting and Managing Opinion Contradictions	1219
<i>Mikalai Tsytsarau and Themis Palpanas</i>	
Demo Papers	
BioGraph: Knowledge Discovery and Exploration in the Biomedical Domain	1223
<i>Jeroen De Knijf, Anthony Liekens, Walter Daelemans, Peter De Rijk, Jurgen Del-Favero, and Bart Goethals</i>	
BeTracker: A System for Finding Behavioral Patterns from Contextual Sensor and Social Data	1227
<i>Hsun-Ping Hsieh, Cheng-Te Li, and Shou-De Lin</i>	

TeamExp: Top- <i>k</i> Team Formation in Social Networks	1231
<i>Mehdi Kargar and Aijun An</i>	
NetDriller: A Powerful Social Network Analysis Tool	1235
<i>Negar Koochakzadeh, Atieh Sarraf, Keivan Kianmehr, Jon Rokne, and Reda Alhaji</i>	
Route Discovery from Mining Uncertain Trajectories	1239
<i>Hechen Liu, Ling-Yin Wei, Yu Zheng, Markus Schneider, and Wen-Chih Peng</i>	
AskUs: An Opinion Search Engine	1243
<i>Shamita Pital, Japinder Singh, and Magdalini Eirinaki</i>	
SentiCorr: Multilingual Sentiment Analysis of Personal Correspondence	1247
<i>Erik Tromp and Mykola Pechenizkiy</i>	

Author Index