

2010 IEEE 10th International Conference on Data Mining (ICDM 2010)

**Sydney, Australia
13-17 December 2010**

Pages 1-608



**IEEE Catalog Number: CFP10278-PRT
ISBN: 978-1-4244-9131-5**

2010 IEEE International Conference on Data Mining

ICDM 2010

Table of Contents

Welcome Message from the Conference Chairs	zvk
Message from the Program Committee Co-Chairs	xviik
Organizing Committee0.....	xx
Program Committee	xxiik

Keynote Abstracts

Mining Billion-node Graphs: Patterns, Generators and Tools	5
<i>Christos Faloutsos</i>	
Assessing the Significance of Groups in High-Dimensional Data	6
<i>Geoff McLachlan</i>	
10 Years of Data Mining Research: Retrospect and Prospect	7
<i>Xindong Wu</i>	

Regular Papers

Detecting Novel Discrepancies in Communication Networks	8
<i>James Abello, Tina Eliassi-Rad, and Nishchal Devanur</i>	
Multi-agent Random Walks for Local Clustering on Graphs.....	18
<i>Morteza Alamgir and Ulrike von Luxburg</i>	
Spatiotemporal Event Detection in Mobility Network	28
<i>Tom S. Au, Rong Duan, Heeyoung Kim, and Guang-Qin Ma</i>	
An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users	38
<i>Tengfei Bao, Hattia Cao, Enhong Chen, Jilei Tian, and Hui Xiong</i>	
Fast and Flexible Multivariate Time Series Subsequence Search	48
<i>Kanishka Bhaduri, Qiang Zhu, Nikunj C. Oza, and Ashok N. Srivastava</i>	
iSAX 2.0: Indexing and Mining One Billion Time Series	58
<i>Alessandro Camerra, Themis Palpanas, Jin Shieh, and Eamonn Keogh</i>	

Abstraction Augmented Markov Models	68
<i>Cornelia Caragea, Adrian Silvescu, Doina Caragea, and Vasant Honavar</i>	
A Graph-Based Approach for Multi-folder Email Classification	78
<i>Sharma Chakravarthy, Aravind Venkatachalam, and Aditya Telang</i>	
Scalable Influence Maximization in Social Networks under the Linear Threshold Model	88
<i>Wei Chen, Yifei Yuan, and Li Zhang</i>	
CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor Networks	98
<i>Alzenny Da Silva, Raja Chiky, and Georges Hebrail</i>	
Bayesian Maximum Margin Clustering	108
<i>Bo Dai, Baogang Hu, and Gang Niu</i>	
Viral Marketing for Multiple Products	118
<i>Samik Datta, Anirban Majumder, and Nisheeth Shrivastava</i>	
Finding Local Anomalies in Very High Dimensional Space	128
<i>Timothy de Vries, Sanjay Chawla, and Michael E. Houle</i>	
PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets	138
<i>Trong Dinh Thac Do, Anne Laurent, and Alexandre Termier</i>	
Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document	148
<i>Lan Du, Wray Lindsay Buntine, and Huidong Jin</i>	
Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach	158
<i>Wouter Duivesteijn, Arno Knobbe, Ad Feelders, and Matthijs van Leeuwen</i>	
Feature Selection for Unsupervised Learning Using Random Cluster Ensembles	168
<i>Haytham Elghazel and Alex Aussem</i>	
Learning Attribute-to-Feature Mappings for Cold-Start Recommendations	176
<i>Zeno Gantner, Lucas Drumond, Christoph Freudenthaler, Steffen Rendle, and Lars Schmidt-Thieme</i>	
An Extensive Empirical Study on Semi-supervised Learning	186
<i>Yuanyuan Guo, Xiaoda Niu, and Harry Zhang</i>	
Efficient Discovery of the Top-K Optimal Dependency Rules with Fisher’s Exact Test of Significance	196
<i>Wilhelmiina Hämmäläinen</i>	
A Variance Reduction Framework for Stable Feature Selection	206
<i>Yue Han and Lei Yu</i>	

Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection	216
<i>Kohei Hayashi, Takashi Takenouchi, Tomohiro Shibata, Yuki Kamiya, Daishi Kato, Kazuo Kunieda, Keiji Yamada, and Kazushi Ikeda</i>	
Rare Category Characterization	226
<i>Jingrui He, Hanghang Tong, and Jaime Carbonell</i>	
Algorithm for Discovering Low-Variance 3-Clusters from Real-Valued Datasets	236
<i>Zhen Hu and Raj Bhatnagar</i>	
Improved Consistent Sampling, Weighted Minhash and L1 Sketching	246
<i>Sergey Ioffe</i>	
An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews	256
<i>Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang</i>	
A Pairwise-Systematic Microaggregation for Statistical Disclosure Control	266
<i>Md. Enamul Kabir, Hua Wang, and Yanchun Zhang</i>	
Multi-label Feature Selection for Graph Classification	274
<i>Xiangnan Kong and Philip S. Yu</i>	
A Binary Decision Diagram-Based One-Class Classifier	284
<i>Takuro Kutsuna</i>	
Detecting Blackhole and Volcano Patterns in Directed Networks	294
<i>Zhongmou Li, Hui Xiong, Yanchi Liu, and Aoying Zhou</i>	
Exploiting Local Data Uncertainty to Boost Global Outlier Detection	304
<i>Bo Liu, Jie Yin, Yanshan Xiao, Longbing Cao, and Philip S. Yu</i>	
Training Conditional Random Fields Using Transfer Learning for Gesture Recognition	314
<i>Jie Liu, Kai Yu, Yi Zhang, and Yalou Huang</i>	
Stratified Sampling for Data Mining on the Deep Web	324
<i>Tantan Liu, Fan Wang, and Gagan Agrawal</i>	
Learning Markov Network Structure with Decision Trees	334
<i>Daniel Lowd and Jesse Davis</i>	
Towards Structural Sparsity: An Explicit l_2/l_0 Approach	344
<i>Dijun Luo, Chris Ding, and Heng Huang</i>	
Multi-document Summarization Using Minimum Distortion	354
<i>Tengfei Ma and Xiaojun Wan</i>	
A Log-Linear Model with Latent Features for Dyadic Prediction	364
<i>Aditya Krishna Menon and Charles Elkan</i>	
Edge Weight Regularization over Multiple Graphs for Similarity Learning	374
<i>Pradeep Muthukrishnan, Dragomir Radev, and Qiaozhu Mei</i>	
A New SVM Approach to Multi-instance Multi-label Learning	384
<i>Nam Nguyen</i>	

Bayesian Aggregation of Binary Classifiers	393
<i>Sunho Park and Seungjin Choi</i>	
Permutations as Angular Data: Efficient Inference in Factorial Spaces	403
<i>Sergey M. Plis, Terran Lane, and Vince D. Calhoun</i>	
Separation of Interleaved Web Sessions with Heuristic Search	411
<i>Marko Poženel, Viljan Mahnic, and Matjaž Kukar</i>	
Consequences of Variability in Classifier Performance Estimates	421
<i>Troy Raeder, T. Ryan Hoens, and Nitesh V. Chawla</i>	
Mining Sensor Streams for Discovering Human Activity Patterns over Time	431
<i>Parisa Rashidi and Diane J. Cook</i>	
Decision Trees for Uplift Modeling	441
<i>Piotr Rzepakowski and Szymon Jaroszewicz</i>	
Co-clustering of Lagged Data	451
<i>Eran Shaham, David Sarne, and Boaz Ben-Moshe</i>	
Polishing the Right Apple: Anytime Classification Also Benefits Data Streams with Constant Arrival Times	461
<i>Jin Shieh and Eamonn Keogh</i>	
Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data	471
<i>Kelvin Sim, Zeyar Aung, and Vivekanand Gopalkrishnan</i>	
gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree Division or Agglomeration	481
<i>Heli Sun, Jianbin Huang, Jiawei Han, Hongbo Deng, Peixiang Zhao, and Boqin Feng</i>	
LogTree: A Framework for Generating System Events from Raw Textual Logs	491
<i>Liang Tang and Tao Li</i>	
Mining Closed Strict Episodes	501
<i>Nikolaj Tatti and Boris Cule</i>	
Multi-dimensional Mass Estimation and Mass-based Clustering	511
<i>Kai Ming Ting and Jonathan R. Wells</i>	
minCENTropy: A Novel Information Theoretic Approach for the Generation of Alternative Clusterings	521
<i>Nguyen Xuan Vinh and Julien Epps</i>	
A Conscience On-line Learning Approach for Kernel-Based Clustering	531
<i>Chang-Dong Wang, Jian-Huang Lai, and Jun-Yong Zhu</i>	
Weighted Feature Subset Non-negative Matrix Factorization and Its Applications to Document Understanding	541
<i>Dingding Wang, Tao Li, and Chris Ding</i>	
Learning a Bi-Stochastic Data Similarity Matrix	551
<i>Fei Wang, Ping Li, and Arnd Christian König</i>	

Active Spectral Clustering	561
<i>Xiang Wang and Ian Davidson</i>	
Discovering Overlapping Groups in Social Media	569
<i>Xufei Wang, Lei Tang, Huiji Gao, and Huan Liu</i>	
Adaptive Distances on Sets of Vectors	579
<i>Adam Woznica and Alexandros Kalousis</i>	
SMILE: A Similarity-Based Approach for Multiple Instance Learning	589
<i>Yanshan Xiao, Bo Liu, Longbing Cao, Jie Yin, and Xindong Wu</i>	
Modeling Information Diffusion in Implicit Networks	599
<i>Jaewon Yang and Jure Leskovec</i>	
Term Filtering with Bounded Error	609
<i>Zi Yang, Wei Li, Jie Tang, and Juanzi Li</i>	
Exploiting Unlabeled Data to Enhance Ensemble Diversity	619
<i>Min-Ling Zhang and Zhi-Hua Zhou</i>	
Constraint Based Dimension Correlation and Distance Divergence for Clustering High-Dimensional Data	629
<i>Xianchao Zhang, Yao Wu, and Yang Qiu</i>	
Active Learning from Multiple Noisy Labelers with Varied Costs	639
<i>Yaling Zheng, Stephen Scott, and Kun Deng</i>	
A Novel Contrast Co-learning Framework for Generating High Quality Training Data	649
<i>Zeyu Zheng, Jun Yan, Shuicheng Yan, Ning Liu, Zheng Chen, and Ming Zhang</i>	
Network Simplification with Minimal Loss of Connectivity	659
<i>Fang Zhou, Sébastien Malher, and Hannu Toivonen</i>	
Improving Kernel Methods through Complex Data Mapping	669
<i>Hang Zhou, Fabio Ramos, and Eric Nettleton</i>	
NESVM: A Fast Gradient Method for Support Vector Machines	679
<i>Tianyi Zhou, Dacheng Tao, and Xindong Wu</i>	
Clustering Large Attributed Graphs: An Efficient Incremental Approach	689
<i>Yang Zhou, Hong Cheng, and Jeffrey Xu Yu</i>	
Mother Fugger: Mining Historical Manuscripts with Local Color Patches	699
<i>Qiang Zhu and Eamonn Keogh</i>	
D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification	709
<i>Fuzhen Zhuang, Ping Luo, Zhiyong Shen, Qing He, Yuhong Xiong, and Zhongzhi Shi</i>	

Short Papers

SONNET: Efficient Approximate Nearest Neighbor Using Multi-core	719
<i>Mohammad Al Hasan, Hilmi Yildirim, and Abhirup Chakraborty</i>	
Two of a Kind or the Ratings Game? Adaptive Pairwise Preferences and Latent Factor Models	725
<i>Suhrid Balakrishnan and Sumit Chopra</i>	
Document Similarity Self-Join with MapReduce	731
<i>Ranieri Baraglia, Gianmarco De Francisci Morales, and Claudio Lucchese</i>	
Quantification via Probability Estimators	737
<i>Antonio Bella, Cèsar Ferri, José Hernández-Orallo, and María José Ramírez-Quintana</i>	
Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks	743
<i>Xiongcai Cai, Michael Bain, Alfred Krzywicki, Wayne Wobcke, Yang Sok Kim, Paul Compton, and Ashesh Mahidadia</i>	
Approximation of Frequentness Probability of Itemsets in Uncertain Data	749
<i>Toon Calders, Calin Garboni, and Bart Goethals</i>	
On Finding Frequent Patterns in Event Sequences	755
<i>Andrea Campagna and Rasmus Pagh</i>	
Active Improvement of Hierarchical Object Features under Budget Constraints	761
<i>Nicolas Cebron</i>	
Pseudo Conditional Random Fields: Joint Training Approach to Segmenting and Labeling Sequence Data	767
<i>Shing-Kit Chan and Wai Lam</i>	
Location and Scatter Matching for Dataset Shift in Text Mining	773
<i>Bo Chen, Wai Lam, Ivor Tsang, and Tak-Lam Wong</i>	
Learning Preferences with Millions of Parameters by Enforcing Sparsity	779
<i>Xi Chen, Bing Bai, Yanjun Qi, Qihang Lin, and Jaime Carbonell</i>	
QMAS: Querying, Mining and Summarization of Multi-modal Databases	785
<i>Robson L.F. Cordeiro, Fan Guo, Donna S. Haverkamp, James H. Horne, Ellen K. Hughes, Gunhee Kim, Agma J.M. Traina, Caetano Traina Jr., and Christos Faloutsos</i>	
Block-GP: Scalable Gaussian Process Regression for Multimodal Data	791
<i>Kamalika Das and Ashok N. Srivastava</i>	
Active Learning with Human-Like Noisy Oracle	797
<i>Jun Du and Charles X. Ling</i>	
Monotone Relabeling in Ordinal Classification	803
<i>Ad Feelders</i>	

The Effect of History on Modeling Systems' Performance: The Problem of the Demanding Lord	809
<i>George Giannakopoulos and Themis Palpanas</i>	
Resilient K-d Trees: K-Means in Space Revisited	815
<i>Fabian Gieseke, Gabriel Moruz, and Jan Vahrenhold</i>	
Advertising Campaigns Management: Should We Be Greedy?	821
<i>Sertan Girgin, Jeremie Mary, Philippe Preux, and Olivier Nicol</i>	
Accelerating Radius-Margin Parameter Selection for SVMs Using Geometric Bounds	827
<i>Ben Goodrich, David Albrecht, and Peter Tischer</i>	
Enhancing Single-Objective Projective Clustering Ensembles	833
<i>Francesco Gullo, Carlotta Domeniconi, and Andrea Tagarelli</i>	
Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects	839
<i>Francesco Gullo, Giovanni Ponti, and Andrea Tagarelli</i>	
Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms	845
<i>Stephan Günnemann, Ines Färber, Brigitte Boden, and Thomas Seidl</i>	
Multi-stream Join Answering for Mining Significant Cross-Stream Correlations	851
<i>Robert Gwadera</i>	
Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service	857
<i>Tsukasa Ishigaki, Takeshi Takenaka, and Yoichi Motomura</i>	
Content-Based Methods for Predicting Web-Site Demographic Attributes	863
<i>Santosh Kabbur, Eui-Hong Han, and George Karypis</i>	
Discrimination Aware Decision Tree Learning	869
<i>Faisal Kamiran, Toon Calders, and Mykola Pechenizkiy</i>	
Patterns on the Connected Components of Terabyte-Scale Graphs	875
<i>U. Kang, Mary McGlohon, Leman Akoglu, and Christos Faloutsos</i>	
Attribution of Conversion Events to Multi-channel Media	881
<i>Brendan Kitts, Liang Wei, Dyng Au, Amanda Powter, and Brian Burdick</i>	
Mining Public Transport Usage for Personalised Intelligent Transport Systems	887
<i>Neal Lathia, Jon Froehlich, and Licia Capra</i>	
Micro-blogging Sentiment Detection by Collaborative Online Learning	893
<i>Guangxia Li, Steven C.H. Hoi, Kuiyu Chang, and Ramesh Jain</i>	
Enforcing Vocabulary k -Anonymity by Semantic Similarity Based Clustering	899
<i>Junqiang Liu and Ke Wang</i>	
Efficient Probabilistic Latent Semantic Analysis with Sparsity Control	905
<i>Sen Liu, Chaolun Xia, and Xiaohong Jiang</i>	
Understanding of Internal Clustering Validation Measures	911
<i>Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, and Junjie Wu</i>	

Transfer Learning via Cluster Correspondence Inference	917
<i>Mingsheng Long, Wei Cheng, Xiaoming Jin, Jianmin Wang, and Dou Shen</i>	
Supervised Link Prediction Using Multiple Sources	923
<i>Zhengdong Lu, Berkant Savas, Wei Tang, and Inderjit S. Dhillon</i>	
Addressing Concept-Evolution in Concept-Drifting Data Streams	929
<i>Mohammad M. Masud, Qing Chen, Latifur Khan, Charu Aggarwal, Jing Gao, Jiawei Han, and Bhavani Thuraisingham</i>	
Sparse Boolean Matrix Factorizations	935
<i>Pauli Miettinen</i>	
On the Computation of Stochastic Search Variable Selection in Linear Regression with UDFs	941
<i>Mario Navas, Carlos Ordonez, and Veerabhadran Baladandayuthapani</i>	
Data Editing Techniques to Allow the Application of Distance-Based Outlier Detection to Streams	947
<i>Vit Niennattrakul, Eamonn Keogh, and Chotirat Ann Ratanamahatana</i>	
Anomaly Detection Using an Ensemble of Feature Models	953
<i>Keith Noto, Carly Brodley, and Donna Slonim</i>	
Assessing Data Mining Results on Matrices with Randomization	959
<i>Markus Ojala</i>	
A Generalized Linear Threshold Model for Multiple Cascades	965
<i>Nishith Pathak, Arindam Banerjee, and Jaideep Srivastava</i>	
Recommending Social Events from Mobile Phone Location Data	971
<i>Daniele Quercia, Neal Lathia, Francesco Calabrese, Giusy Di Lorenzo, and Jon Crowcroft</i>	
On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index	977
<i>R. Quèrè, H. Le Capitaine, N. Fraisseix, and C. Frélicot</i>	
Financial Forecasting with Gompertz Multiple Kernel Learning	983
<i>Han Qin, Dejing Dou, and Yue Fang</i>	
Leveraging D-Separation for Relational Data Sets	989
<i>Matthew J.H. Rattigan and David Jensen</i>	
Factorization Machines	995
<i>Steffen Rendle</i>	
Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs	1001
<i>Doruk Sart, Abdullah Mueen, Walid Najjar, Eamonn Keogh, and Vit Niennattrakul</i>	
An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection	1007
<i>Tim Schlüter and Stefan Conrad</i>	
Bonsai: Growing Interesting Small Trees	1013
<i>Stephan Seufert, Srikantha Bedathur, Julian Mestre, and Gerhard Weikum</i>	

Mixed-Membership Stochastic Block-Models for Transactional Networks	1019
<i>Mahdi Shafiei and Hugh Chipman</i>	
Generalized Probabilistic Matrix Factorizations for Collaborative Filtering	1025
<i>Hanhuai Shan and Arindam Banerjee</i>	
Topic Modeling Ensembles	1031
<i>Zhiyong Shen, Ping Luo, Shengwen Yang, and Xukun Shen</i>	
Interval-valued Matrix Factorization with Applications	1037
<i>Zhiyong Shen, Liang Du, Xukun Shen, and Yidong Shen</i>	
Efficient Semi-supervised Spectral Co-clustering with Constraints	1043
<i>Xiaoxiao Shi, Wei Fan, and Philip S. Yu</i>	
Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation	1049
<i>Xiaoxiao Shi, Qi Liu, Wei Fan, Philip S. Yu, and Ruixin Zhu</i>	
One-Class Matrix Completion with Low-Density Factorizations	1055
<i>Vikas Sindhwani, Serhat S. Bucak, Jianying Hu, and Aleksandra Mojsilovic</i>	
A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support	1061
<i>Jimeng Sun, Daby Sow, Jianying Hu, and Shahram Ebadollahi</i>	
Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust, and Fast Training Method	1067
<i>Xu Sun, Hisashi Kashima, Takuya Matsuzaki, and Naonori Ueda</i>	
Visualizing Graphs Using Minimum Spanning Dendrograms	1073
<i>Daniel Svonava and Michail Vlachos</i>	
Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems	1079
<i>Lu-An Tang, Xiao Yu, Sangkyum Kim, Jiawei Han, Chih-Chieh Hung, and Wen-Chih Peng</i>	
Node Similarities from Spreading Activation	1085
<i>Kilian Thiel and Michael R. Berthold</i>	
On the Vulnerability of Large Graphs	1091
<i>Hanghang Tong, B. Aditya Prakash, Charalampos Tsourakakis, Tina Eliassi-Rad, Christos Faloutsos, and Duen Horng Chau</i>	
Testing the Significance of Patterns in Data with Cluster Structure	1097
<i>Niko Vuokko and Petteri Kaski</i>	
Compressed Nonnegative Sparse Coding	1103
<i>Fei Wang and Ping Li</i>	
Anonymizing Temporal Data	1109
<i>Ke Wang, Yabo Xu, Raymond Chi-Wing Wong, and Ada Wai-Chee Fu</i>	
Homotopy Regularization for Boosting	1115
<i>Zheng Wang, Yangqiu Song, and Changshui Zhang</i>	

What Do People Want in Microblogs? Measuring Interestingness of Hashtags in <i>Twitter</i>	1121
<i>Jianshu Weng, Ee-Peng Lim, Qi He, and Cane Wing-Ki Leung</i>	
Probabilistic Inference Protection on Anonymized Data	1127
<i>Raymond Chi-Wing Wong, Ada Wai-Chee Fu, Ke Wang, Yabo Xu, Jian Pei, and Philip S. Yu</i>	
Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval	1133
<i>Jun Wu, Ming-Yu Lu, and Chun-Li Wang</i>	
Max-Clique: A Top-Down Graph-Based Approach to Frequent Pattern Mining	1139
<i>Yan Xie and Philip S. Yu</i>	
Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model	1145
<i>Qingyan Yang, Ju Fan, Jianyong Wang, and Lizhu Zhou</i>	
Passive Sampling for Regression	1151
<i>Hwanjo Yu and Sungchul Kim</i>	
Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling	1157
<i>Jun Yu, Weng-Keen Wong, and Rebecca A. Hutchinson</i>	
Causal Discovery from Streaming Features	1163
<i>Kui Yu, Xindong Wu, Hao Wang, and Wei Ding</i>	
ABS: The Anti Bouncing Model for Usage Data Streams	1169
<i>Chongsheng Zhang, Florent Masseglia, and Yves Lechevallier</i>	
Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams	1175
<i>Peng Zhang, Xingquan Zhu, Jianlong Tan, and Li Guo</i>	
Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation	1181
<i>Xianchao Zhang, Yansheng Jiang, Wenxin Liang, and Xin Han</i>	
K-AP: Generating Specified <i>K</i> Clusters by Efficient Affinity Propagation	1187
<i>Xiangliang Zhang, Wei Wang, Kjetil Nørnvåg, and Michèle Sebag</i>	
MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model	1193
<i>Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao</i>	
Hierarchical Ensemble Clustering	1199
<i>Li Zheng, Tao Li, and Chris Ding</i>	
Frequent Instruction Sequential Pattern Mining in Hardware Sample Data	1205
<i>Jia Zou, Jing Xiao, Rui Hou, and Yanqi Wang</i>	
Efficient Episode Mining with Minimal and Non-overlapping Occurrences	1211
<i>Huisheng Zhu, Peng Wang, Xianmang He, Yujia Li, Wei Wang, and Baile Shi</i>	

Tutorials

Spatial and Spatio-temporal Data Mining	1217
<i>Vania Bogorny and Shashi Shekhar</i>	
Knowledge Discovery in Academic Drug Discovery Programs: Opportunities and Challenges	1218
<i>Jun Huan</i>	
How to Do Good Data Mining Research and Get it Published in Top Venues	1219
<i>Eamonn Keogh</i>	
Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data	1220
<i>Emmanuel Müller, Stephan Günemann, Ines Färber, and Thomas Seidl</i>	

Author Index