

2016 IEEE 16th International Conference on Data Mining (ICDM 2016)

**Barcelona, Spain
12 - 15 December 2016**

Pages 1-688



**IEEE Catalog Number: CFP16278-POD
ISBN: 978-1-5090-5474-9**

**Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP16278-POD
ISBN (Print-On-Demand):	978-1-5090-5474-9
ISBN (Online):	978-1-5090-5473-2
ISSN:	1550-4786

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2016 IEEE 16th International Conference on Data Mining

ICDM 2016

Table of Contents

Message from the Conference General Chairs	xvii
Message from the Program Co-Chairs	xix
Organizing Committee	xxi
Program Committee	xxiii
Keynotes	xxxi

Regular Papers

Auditing Black-Box Models for Indirect Influence	1
<i>Philip Adler, Casey Falk, Sorelle A. Friedler, Gabriel Rybeck, Carlos Scheidegger, Brandon Smith, and Suresh Venkatasubramanian</i>	
Asynchronous Multi-task Learning	11
<i>Inci M. Baytas, Ming Yan, Anil K. Jain, and Jiayu Zhou</i>	
Unsupervised Exceptional Attributed Sub-Graph Mining in Urban Data	21
<i>Ahmed Anes Bendimerad, Marc Plantevit, and Céline Robardet</i>	
ADAGIO: Fast Data-Aware Near-Isometric Linear Embeddings	31
<i>Jaroslav Blasiok and Charalampos E. Tsourakakis</i>	
Causal Inference by Compression	41
<i>Kailash Budhathoki and Jilles Vreeken</i>	
On Dense Subgraphs in Signed Network Streams	51
<i>Jose Cadena, Anil Kumar Vullikanti, and Charu C. Aggarwal</i>	
Relief of Spatiotemporal Accessibility Overloading with Optimal Resource Placement	61
<i>Chien-Wei Chang, Hao-Yi Chih, Dean Chou, Yu-Chen Shu, and Kun-Ta Chuang</i>	
Mining Graphlet Counts in Online Social Networks	71
<i>Xiaowei Chen and John C. S. Lui</i>	
Differentially Private Regression Diagnostics	81
<i>Yan Chen, Ashwin Machanavajhala, Jerome P. Reiter, and Andrés F. Barrientos</i>	

Vote-and-Comment: Modeling the Coevolution of User Interactions in Social Voting	
Web Sites	91
<i>Alceu Ferraz Costa, Agma Juci Machado Traina, Caetano Traina Jr., and Christos Faloutsos</i>	
On Efficient External-Memory Triangle Listing	101
<i>Yi Cui, Di Xiao, and Dmitri Loguinov</i>	
Efficient Distributed SGD with Variance Reduction	111
<i>Soham De and Tom Goldstein</i>	
Triply Stochastic Variational Inference for Non-linear Beta Process Factor Analysis	121
<i>Kai Fan, Yizhe Zhang, Ricardo Henao, and Katherine Heller</i>	
Beyond Points and Paths: Counting Private Bodies	131
<i>Maryam Fanaeepour and Benjamin I. P. Rubinstein</i>	
Efficient Rectangular Maximal-Volume Algorithm for Rating Elicitation in Collaborative Filtering	141
<i>Alexander Fonarev, Alexander Mikhalev, Pavel Serdyukov, Gleb Gusev, and Ivan Oseledets</i>	
New Robust Clustering Model for Identifying Cancer Genome Landscapes	151
<i>Hongchang Gao, Xiaoqian Wang, and Heng Huang</i>	
Event Series Prediction via Non-Homogeneous Poisson Process Modelling	161
<i>James Goulding, Simon Preston, and Gavin Smith</i>	
Model Accuracy and Runtime Tradeoff in Distributed Deep Learning: A Systematic Study	171
<i>Suyog Gupta, Wei Zhang, and Fei Wang</i>	
Waddling Random Walk: Fast and Accurate Mining of Motif Statistics in Large Graphs	181
<i>Guyue Han and Harish Sethu</i>	
Fusing Similarity Models with Markov Chains for Sparse Sequential Recommendation	191
<i>Ruining He and Julian McAuley</i>	
Adaptive Neighborhood Propagation by Joint L _{2,1} -Norm Regularized Sparse Coding for Representation and Classification	201
<i>Lei Jia, Zhao Zhang, Lei Wang, Weiming Jiang, and Mingbo Zhao</i>	
From Sets of Good Redescriptions to Good Sets of Redescriptions	211
<i>Janis Kalofolias, Esther Galbrun, and Pauli Miettinen</i>	
Edge Weight Prediction in Weighted Signed Networks	221
<i>Srijan Kumar, Francesca Spezzano, V. S. Subrahmanian, and Christos Faloutsos</i>	
Transfer Learning for Survival Analysis via Efficient L _{2,1} -Norm Regularized Cox Regression	231
<i>Yan Li, Lu Wang, Jie Wang, Jieping Ye, and Chandan K. Reddy</i>	
Interactive Multi-task Relationship Learning	241
<i>Kaixiang Lin and Jiayu Zhou</i>	

Augmented LSTM Framework to Construct Medical Self-Diagnosis Android	251
<i>Chaochun Liu, Huan Sun, Nan Du, Shulong Tan, Hongliang Fei, Wei Fan, Tao Yang, Hao Wu, Yaliang Li, and Chenwei Zhang</i>	
The Optimal Distribution of Electric-Vehicle Chargers across a City	261
<i>Chen Liu, Ke Deng, Chaojie Li, Jianxin Li, Yanhua Li, and Jun Luo</i>	
Streaming Model Selection via Online Factorized Asymptotic Bayesian Inference	271
<i>Chunchen Liu, Lu Feng, and Ryohei Fujimaki</i>	
Robust Multi-View Feature Selection	281
<i>Hongfu Liu, Haiyi Mao, and Yun Fu</i>	
KNN Classifier with Self Adjusting Memory for Heterogeneous Concept Drift	291
<i>Viktor Losing, Barbara Hammer, and Heiko Wersing</i>	
New Probabilistic Multi-graph Decomposition Model to Identify Consistent Human Brain Network Modules	301
<i>Dijun Luo, Zhouyuan Huo, Yang Wang, Andrew J. Saykin, Li Shen, and Heng Huang</i>	
Efficient Extraction of Non-negative Latent Factors from High-Dimensional and Sparse Matrices in Industrial Applications	311
<i>Xin Luo, Mingsheng Shang, and Shuai Li</i>	
Robust Graph-Theoretic Clustering Approaches Using Node-Based Resilience Measures	320
<i>John Matta, Tayo Obafemi-Ajayi, Jeffrey Borwey, Donald Wunsch, and Gunes Ercal</i>	
Hyperbolae are No Hyperbole: Modelling Communities That are Not Cliques	330
<i>Saskia Metzler, Stephan Günnemann, and Pauli Miettinen</i>	
Learning Hierarchically Decomposable Concepts with Active Over-Labeling	340
<i>Yuji Mo, Stephen D. Scott, and Doug Downey</i>	
AWarp: Fast Warping Distance for Sparse Time Series	350
<i>Abdullah Mueen, Nikan Chavoshi, Noor Abu-El-Rub, Hossein Hamooni, and Amanda Minnich</i>	
Binary Classifier Calibration Using an Ensemble of Near Isotonic Regression Models	360
<i>Mahdi Pakdaman Naeini and Gregory F. Cooper</i>	
A Scalable and Generic Framework to Mine Top-k Representative Subgraph Patterns	370
<i>Dheepikaa Natarajan and Sayan Ranu</i>	
What You Will Gain By Rounding: Theory and Algorithms for Rounding Rank	380
<i>Stefan Neumann, Rainer Gemulla, and Pauli Miettinen</i>	
A Fast Iterative Algorithm for Improved Unsupervised Feature Selection	390
<i>Bruno Ordozgoiti, Sandra Gómez Canaval, and Alberto Mozo</i>	
Sparse Factorization Machines for Click-through Rate Prediction	400
<i>Zhen Pan, Enhong Chen, Qi Liu, Tong Xu, Haiping Ma, and Hongjie Lin</i>	

Unsupervised Feature Selection for Outlier Detection by Modelling Hierarchical Value-Feature Couplings	410
<i>Guansong Pang, Longbing Cao, Ling Chen, and Huan Liu</i>	
Partition Aware Connected Component Computation in Distributed Systems	420
<i>Ha-Myung Park, Namyong Park, Sung-Hyon Myaeng, and U. Kang</i>	
Iteratively Reweighted Least Squares Algorithms for L1-Norm Principal Component Analysis	430
<i>Young Woong Park and Diego Klabjan</i>	
Convolutional MKL Based Multimodal Emotion Recognition and Sentiment Analysis	439
<i>Soujanya Poria, Iti Chaturvedi, Erik Cambria, and Amir Hussain</i>	
Recommending Packages to Groups	449
<i>Shuyao Qi, Nikos Mamoulis, Evaggelia Pitoura, and Panayiotis Tsaparas</i>	
Subspace Outlier Detection in Linear Time with Randomized Hashing	459
<i>Saket Sathe and Charu C. Aggarwal</i>	
CoreScope: Graph Mining Using k-Core Analysis — Patterns, Anomalies and Algorithms	469
<i>Kijung Shin, Tina Eliassi-Rad, and Christos Faloutsos</i>	
L-EnsNMF: Boosted Local Topic Discovery via Ensemble of Nonnegative Matrix Factorization	479
<i>Sangho Suh, Jaegul Choo, Joonseok Lee, and Chandan K. Reddy</i>	
Modeling Ambiguity, Subjectivity, and Diverging Viewpoints in Opinion Question Answering Systems	489
<i>Mengting Wan and Julian McAuley</i>	
Traffic Speed Prediction and Congestion Source Exploration: A Deep Learning Method	499
<i>Jingyuan Wang, Qian Gu, Junjie Wu, Guannan Liu, and Zhang Xiong</i>	
Blind Men and The Elephant: Thurstonian Pairwise Preference for Ranking in Crowdsourcing	509
<i>Xiaolong Wang, Jingjing Wang, Luo Jie, Chengxiang Zhai, and Yi Chang</i>	
Regularizing Deep Convolutional Neural Networks with a Structured Decorrelation Constraint	519
<i>Wei Xiong, Bo Du, Lefei Zhang, Ruimin Hu, and Dacheng Tao</i>	
Aligned Matrix Completion: Integrating Consistency and Independency in Multiple Domains	529
<i>Linli Xu, Zaiyi Chen, Qi Zhou, Enhong Chen, Nicholas Jing Yuan, and Xing Xie</i>	
Heterogeneous Representation Learning with Structured Sparsity Regularization	539
<i>Pei Yang and Jingrui He</i>	

POI Recommendation: A Temporal Matching between POI Popularity and User Regularity	549
<i>Zijun Yao, Yanjie Fu, Bin Liu, Yanchi Liu, and Hui Xiong</i>	
College Student Scholarships and Subsidies Granting: A Multi-modal Multi-label Approach	559
<i>Han-Jia Ye, De-Chuan Zhan, Xiaolin Li, Zhen-Chuan Huang, and Yuan Jiang</i>	
Generalized Independent Subspace Clustering	569
<i>Wei Ye, Samuel Maurus, Nina Hubig, and Claudia Plant</i>	
Matrix Profile III: The Matrix Profile Allows Visualization of Salient Subsequences in Massive Time Series	579
<i>Chin-Chia Michael Yeh, Helga Van Herle, and Eamonn Keogh</i>	
Random Walk with Restart over Dynamic Graphs	589
<i>Weiren Yu and Julie McCann</i>	
Communities in Preference Networks: Refined Axioms and Beyond	599
<i>Gang Zeng, Yuyi Wang, Juhua Pu, Xingwu Liu, Xiaoming Sun, and Jialin Zhang</i>	
Homophily, Structure, and Content Augmented Network Representation Learning	609
<i>Daokun Zhang, Jie Yin, Xingquan Zhu, and Chengqi Zhang</i>	
Fixing the Convergence Problems in Parallel Asynchronous Dual Coordinate Descent	619
<i>Huan Zhang and Cho-Jui Hsieh</i>	
HogWild++: A New Mechanism for Decentralized Asynchronous Stochastic Gradient Descent	629
<i>Huan Zhang, Cho-Jui Hsieh, and Venkatesh Akella</i>	
Predicting COPD Failure by Modeling Hazard in Longitudinal Clinical Data	639
<i>Jianfei Zhang, Shengrui Wang, Josiane Courteau, Lifei Chen, Aurélien Bach, and Alain Vanasse</i>	
Reliable Gender Prediction Based on Users' Video Viewing Behavior	649
<i>Jie Zhang, Kuang Du, Ruihua Cheng, Zhi Wei, Chenguang Qin, Huaxin You, and Sha Hu</i>	
Probabilistic-Mismatch Anomaly Detection: Do One's Medications Match with the Diagnoses	659
<i>Lingxiao Zhang, Xiang Li, Haifeng Liu, Jing Mei, Gang Hu, Junfeng Zhao, Yanzhen Zou, Bing Xie, and Guotong Xie</i>	
Inferring Latent Network from Cascade Data for Dynamic Social Recommendation	669
<i>Qin Zhang, Jia Wu, Peng Zhang, Guodong Long, Ivor W. Tsang, and Chengqi Zhang</i>	
Group Preference Aggregation: A Nash Equilibrium Approach	679
<i>Hongke Zhao, Qi Liu, Yong Ge, Ruoyan Kong, and Enhong Chen</i>	
Multi-resolution Spatial Event Forecasting in Social Media	689
<i>Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan</i>	
To be or Not to be Friends: Exploiting Social Ties for Venture Investments	699
<i>Hao Zhong, Chuanren Liu, Xinjiang Lu, and Hui Xiong</i>	

Graph-Structured Sparse Optimization for Connected Subgraph Detection	709
<i>Baojian Zhou and Feng Chen</i>	
Bi-Level Rare Temporal Pattern Detection	719
<i>Dawei Zhou, Jingrui He, Yu Cao, and Jae-Sun Seo</i>	
A Bayesian Nonparametric Approach to Dynamic Dyadic Data Prediction	729
<i>Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng</i>	
Matrix Profile II: Exploiting a Novel Algorithm and GPUs to Break the One Hundred Million Barrier for Time Series Motifs and Joins	739
<i>Yan Zhu, Zachary Zimmerman, Nader Shakibay Senobari, Chin-Chia Michael Yeh, Gareth Funning, Abdullah Mueen, Philip Brisk, and Eamonn Keogh</i>	
Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding	749
<i>Zihao Zhu, Changchang Yin, Buyue Qian, Yu Cheng, Jishang Wei, and Fei Wang</i>	
ConTrack: A Scalable Method for Tracking Multiple Concepts in Large Scale Multidimensional Data	759
<i>Ali Zonoozi, Qirong Ho, Shonali Krishnaswamy, and Gao Cong</i>	

Short Papers

Fractality of Massive Graphs: Scalable Analysis with Sketch-Based Box-Covering Algorithm	769
<i>Takuya Akiba, Kenko Nakamura, and Taro Takaguchi</i>	
Cut Tree Construction from Massive Graphs	775
<i>Takuya Akiba, Yoichi Iwata, Yosuke Sameshima, Naoto Mizuno, and Yosuke Yano</i>	
Learning from Your Network of Friends: A Trajectory Representation Learning Model Based on Online Social Ties	781
<i>Basma Alharbi and Xiangliang Zhang</i>	
A Combinatorial Approach to Role Discovery	787
<i>Albert Arockiasamy, Aristides Gionis, and Nikolaj Tatti</i>	
DESQ: Frequent Sequence Mining with Subsequence Constraints	793
<i>Kaustubh Beedkar and Rainer Gemulla</i>	
EXTRACT: Strong Examples from Weakly-Labeled Sensor Data	799
<i>Davis W. Blalock and John V. Guttag</i>	
A Theoretical Analysis of the Fuzzy K-Means Problem	805
<i>Johannes Blömer, Sascha Brauer, and Kathrin Bujna</i>	
Efficient Sampling-Based Kernel Mean Matching	811
<i>Swarup Chandra, Ahsanul Haque, Latifur Khan, and Charu Aggarwal</i>	
DeBot: Twitter Bot Detection via Warped Correlation	817
<i>Nikan Chavoshi, Hossein Hamooni, and Abdullah Mueen</i>	

Interpretable Clustering via Discriminative Rectangle Mixture Model	823
<i>Junxiang Chen, Yale Chang, Brian Hobbs, Peter Castaldi, Michael Cho, Edwin Silverman, and Jennifer Dy</i>	
Asymptotic Analysis of Equivalences and Core-Structures in Kronecker-Style Graph Models	829
<i>Alex J. Chin, Timothy D. Goodrich, Michael P. O'Brien, Felix Reidl, Blair D. Sullivan, and Andrew van der Poel</i>	
Event Grounding from Multimodal Social Network Fusion	835
<i>Hyunsouk Cho, Jinyoung Yeo, and Seung-Won Hwang</i>	
Can Active Learning Experience Be Transferred?	841
<i>Hong-Min Chu and Hsuan-Tien Lin</i>	
Outlier Detection from Network Data with Subnetwork Interpretation	847
<i>Xuan-Hong Dang, Arlei Silva, Ambuj Singh, Ananthram Swami, and Prithwish Basu</i>	
Incorporating Expert Feedback into Active Anomaly Discovery	853
<i>Shubhomoy Das, Weng-Keen Wong, Thomas Dietterich, Alan Fern, and Andrew Emmott</i>	
Spell: Streaming Parsing of System Event Logs	859
<i>Min Du and Feifei Li</i>	
Modeling Time Lags in Citation Networks	865
<i>Tao-Yang Fu, Zhen Lei, and Wang-Chien Lee</i>	
Efficient and Distributed Algorithms for Large-Scale Generalized Canonical Correlations Analysis	871
<i>Xiao Fu, Kejun Huang, Evangelos E. Papalexakis, Hyun-Ah Song, Partha Pratim Talukdar, Nicholas D. Sidiropoulos, Christos Faloutsos, and Tom Mitchell</i>	
Service Usage Analysis in Mobile Messaging Apps: A Multi-label Multi-view Perspective	877
<i>Yanjie Fu, Junming Liu, Xiaolin Li, Xinjiang Lu, Jingci Ming, Chu Guan, and Hui Xiong</i>	
A Semi-Supervised AUC Optimization Method with Generative Models	883
<i>Akinori Fujino and Naonori Ueda</i>	
MeGS: Partitioning Meaningful Subgraph Structures Using Minimum Description Length	889
<i>Sebastian Goebel, Annika Tonch, Christian Böhm, and Claudia Plant</i>	
Probabilistic Formulations of Regression with Mixed Guidance	895
<i>Aubrey Gress and Ian Davidson</i>	
HLGPS: A Home Location Global Positioning System in Location-Based Social Networks	901
<i>Yulong Gu, Jiaying Song, Weidong Liu, and Lixin Zou</i>	
Large-Scale Embedding Learning in Heterogeneous Event Data	907
<i>Huan Gui, Jialu Liu, Fangbo Tao, Meng Jiang, Brandon Norick, and Jiawei Han</i>	

Direct Mining of Subjectively Interesting Relational Patterns	913
<i>Tias Guns, Achille Aknin, Jeffrey Lijffijt, and Tijl De Bie</i>	
Semi-Supervised Multi-label Dimensionality Reduction	919
<i>Baolin Guo, Chenping Hou, Feiping Nie, and Dongyun Yi</i>	
A Novel Uncertainty Sampling Algorithm for Cost-Sensitive Multiclass Active Learning	925
<i>Kuan-Hao Huang and Hsuan-Tien Lin</i>	
SOAL: Second-Order Online Active Learning	931
<i>Shuji Hao, Peilin Zhao, Jing Lu, Steven C. H. Hoi, Chunyan Miao, and Chi Zhang</i>	
Learning Compatibility Across Categories for Heterogeneous Item Recommendation	937
<i>Ruining He, Charles Packer, and Julian McAuley</i>	
HNP3: A Hierarchical Nonparametric Point Process for Modeling Content Diffusion over Social Media	943
<i>Seyed Abbas Hosseini, Ali Khodadadi, Soheil Arabzade, and Hamid R. Rabiee</i>	
Improved and Scalable Bradley-Terry Model for Collaborative Ranking	949
<i>Jun Hu and Ping Li</i>	
Sparse Gaussian Markov Random Field Mixtures for Anomaly Detection	955
<i>Tsuyoshi Idé, Ankush Khandehwal, and Jayant Kalagnanam</i>	
A Robust Framework for Classifying Evolving Document Streams in an Expert-Machine-Crowd Setting	961
<i>Muhammad Imran, Sanjay Chawla, and Carlos Castillo</i>	
Learning Deep Networks from Noisy Labels with Dropout Regularization	967
<i>Ishan Jindal, Matthew Nokleby, and Xuewen Chen</i>	
Personalized Ranking in Signed Networks Using Signed Random Walk with Restart	973
<i>Jinhong Jung, Woojeong Jin, Lee Sael, and U Kang</i>	
ExploreKit: Automatic Feature Generation and Selection	979
<i>Gilad Katz, Eui Chul Richard Shin, and Dawn Song</i>	
Steering Social Media Promotions with Effective Strategies	985
<i>Kun Kuang, Meng Jiang, Peng Cui, and Shiqiang Yang</i>	
Mining Statistically Significant Attribute Associations in Attributed Graphs	991
<i>Jihwan Lee, Keehwan Park, and Sunil Prabhakar</i>	
Time-Aware User Identification with Topic Models	997
<i>Clément Lesaege, François Schnitzler, Anne Lambert, and Jean-Ronan Vigouroux</i>	
Toward Time-Evolving Feature Selection on Dynamic Networks	1003
<i>Jundong Li, Xia Hu, Ling Jian, and Huan Liu</i>	
Concept Based Short Text Stream Classification with Topic Drifting Detection	1009
<i>Peipei Li, Lu He, Xuegang Hu, Yuhong Zhang, Lei Li, and Xindong Wu</i>	

Regularized Large Margin Distance Metric Learning	1015
<i>Ya Li, Xinmei Tian, and Dacheng Tao</i>	
Mutual Reinforcement of Academic Performance Prediction and Library Book Recommendation	1023
<i>Defu Lian, Yuyang Ye, Wenya Zhu, Qi Liu, Xing Xie, and Hui Xiong</i>	
Regularized Content-Aware Tensor Factorization Meets Temporal-Aware Location Recommendation	1029
<i>Defu Lian, Zhenyu Zhang, Yong Ge, Fuzheng Zhang, Nicholas Jing Yuan, and Xing Xie</i>	
Whether This Participant will Attract You to This Event? Exploiting Participant Influence for Event Recommendation	1035
<i>Yi Liao, Xinshi Lin, and Wai Lam</i>	
HIVE-COTE: The Hierarchical Vote Collective of Transformation-Based Ensembles for Time Series Classification	1041
<i>Jason Lines, Sarah Taylor, and Anthony Bagnall</i>	
House Price Modeling over Heterogeneous Regions with Hierarchical Spatial Functional Analysis	1047
<i>Bang Liu, Borislav Mavrin, Di Niu, and Linglong Kong</i>	
Context-Aware Sequential Recommendation	1053
<i>Qiang Liu, Shu Wu, Diyi Wang, Zhaokang Li, and Liang Wang</i>	
Structure-Preserved Multi-source Domain Adaptation	1059
<i>Hongfu Liu, Ming Shao, and Yun Fu</i>	
Sublinear Dual Coordinate Ascent for Regularized Loss Minimization	1065
<i>Liu Liu and Dacheng Tao</i>	
Efficient and Scalable Detection of Overlapping Communities in Big Networks	1071
<i>Tianshu Lyu, Lidong Bing, Zhao Zhang, and Yan Zhang</i>	
Factorizing Complex Discrete Data “with Finesse”	1077
<i>Samuel Maurus and Claudia Plant</i>	
Towards Scalable Network Delay Minimization	1083
<i>Sourav Medya, Petko Bogdanov, and Ambuj Singh</i>	
Foundations of Perturbation Robust Clustering	1089
<i>Jarrod Moore and Margareta Ackerman</i>	
Faster Kernels for Graphs with Continuous Attributes via Hashing	1095
<i>Christopher Morris, Nils M. Kriege, Kristian Kersting, and Petra Mutzel</i>	
Optimizing the Multiclass F-Measure via Biconcave Programming	1101
<i>Harikrishna Narasimhan, Weiwei Pan, Purushottam Kar, Pavlos Protopapas, and Harish G. Ramaswamy</i>	
Budgeted Batch Bayesian Optimization	1107
<i>Vu Nguyen, Santu Rana, Sunil K Gupta, Cheng Li, and Svetha Venkatesh</i>	

One-Pass Logistic Regression for Label-Drift and Large-Scale Classification on Distributed Systems	1113
<i>Vu Nguyen, Tu Dinh Nguyen, Trung Le, Svetha Venkatesh, and Dinh Phung</i>	
Self-Grouping Multi-network Clustering	1119
<i>Jingchao Ni, Wei Cheng, Wei Fan, and Xiang Zhang</i>	
A Scalable Framework for Stylometric Analysis Query Processing	1125
<i>Sarana Nutanong, Chenyun Yu, Raheem Sarwar, Peter Xu, and Dickson Chow</i>	
Compressing Random Forests	1131
<i>Amichai Painsky and Saharon Rosset</i>	
A Fast Factorization-Based Approach to Robust PCA	1137
<i>Chong Peng, Zhao Kang, and Qiang Cheng</i>	
Background Check: A General Technique to Build More Reliable and Versatile Classifiers	1143
<i>Miquel Perello-Nieto, Telmo De Menezes E Silva Filho, Meelis Kull, and Peter Flach</i>	
Product-Based Neural Networks for User Response Prediction	1149
<i>Yanru Qu, Han Cai, Kan Ren, Weinan Zhang, Yong Yu, Ying Wen, and Jun Wang</i>	
Efficient Algorithms for the Three Locus Problem in Genome-Wide Association Study	1155
<i>Sanguthevar Rajasekaran and Subrata Saha</i>	
Scalable Block Scheduling for Efficient Multi-database Record Linkage	1161
<i>Thilina Ranbaduge, Dinusha Vatsalan, and Peter Christen</i>	
Sequential Ensemble Learning for Outlier Detection: A Bias-Variance Perspective	1167
<i>Shebuti Rayana, Wen Zhong, and Leman Akoglu</i>	
Learning Independent, Diverse Binary Hash Functions: Pruning and Locality	1173
<i>Ramin Raziperchikolaei and Miguel Á. Carreira-Perpiñán</i>	
A Rotation Invariant Latent Factor Model for Moveme Discovery from Static Poses	1179
<i>Matteo Ruggero Ronchi, Joon Sik Kim, and Yisong Yue</i>	
Patterns in Cognitive Rehabilitation of Traumatic Brain Injury Patients: A Text Mining Approach	1185
<i>Alejandro García Rudolph, Alberto García Molina, Eloy Opisso, and Josep María Tormos</i>	
Detecting Change Processes in Dynamic Networks by Frequent Graph Evolution Rule Mining	1191
<i>Erik Scharwächter, Emmanuel Müller, Jonathan Donges, Marwan Hassani, and Thomas Seidl</i>	
Reliable Semi-supervised Learning	1197
<i>Junming Shao, Chen Huang, Qinli Yang, and Guangchun Luo</i>	
Online Unsupervised Multi-view Feature Selection	1203
<i>Weixiang Shao, Lifang He, Chun-Ta Lu, Xiaokai Wei, and Philip S. Yu</i>	

Prefix and Suffix Invariant Dynamic Time Warping	1209
<i>Diego F. Silva, Gustavo E. A. P. A. Batista, and Eamonn Keogh</i>	
Mining Summaries for Knowledge Graph Search	1215
<i>Qi Song, Yinghui Wu, and Xin Luna Dong</i>	
Structure Selection for Convolutional Non-negative Matrix Factorization Using Normalized Maximum Likelihood Coding	1221
<i>Atsushi Suzuki, Kohei Miyaguchi, and Kenji Yamanishi</i>	
Modeling Real Estate for School District Identification	1227
<i>Fei Tan, Chaoran Cheng, and Zhi Wei</i>	
Feature Grouping Using Weighted ℓ_1 Norm for High-Dimensional Data	1233
<i>Bhanukiran Vinzamuri, Karthik K. Padthe, and Chandan K. Reddy</i>	
Learning Task Relational Structure for Multi-task Feature Learning	1239
<i>De Wang, Feiping Nie, and Heng Huang</i>	
Multi-view Clustering via Concept Factorization with Local Manifold Regularization	1245
<i>Hao Wang, Yan Yang, and Tianrui Li</i>	
A Fast Distributed Classification Algorithm for Large-Scale Imbalanced Data	1251
<i>Huihui Wang, Yang Gao, Yinghuan Shi, and Hao Wang</i>	
Differential Location Privacy for Sparse Mobile Crowdsensing	1257
<i>Leye Wang, Daqing Zhang, Dingqi Yang, Brian Y. Lim, and Xiaojuan Ma</i>	
Robust Convex Clustering Analysis	1263
<i>Qi Wang, Pinghua Gong, Shiyu Chang, Thomas S. Huang, and Jiayu Zhou</i>	
Bayesian Rule Sets for Interpretable Classification	1269
<i>Tong Wang, Cynthia Rudin, Finale Velez-Doshi, Yimin Liu, Erica Klampfl, and Perry MacNeille</i>	
The Development of a Smart Taxicab Scheduling System: A Multi-source Data Fusion Perspective	1275
<i>Yang Wang, Binxin Liang, Wei Zheng, Liusheng Huang, and Hengchang Liu</i>	
Selecting Valuable Customers for Merchants in E-Commerce Platforms	1281
<i>Yijun Wang, Le Wu, Zongda Wu, Enhong Chen, and Qi Liu</i>	
Canonical Consistent Weighted Sampling for Real-Value Weighted Min-Hash	1287
<i>Wei Wu, Bin Li, Ling Chen, and Chengqi Zhang</i>	
Spectrum-Revealing Cholesky Factorization for Kernel Methods	1293
<i>Jianwei Xiao and Ming Gu</i>	
Topic Discovery for Short Texts Using Word Embeddings	1299
<i>Guangxu Xun, Vishrawas Gopalakrishnan, Fenglong Ma, Yaliang Li, Jing Gao, and Aidong Zhang</i>	
Dynamic Contextual Multi Arm Bandits in Display Advertisement	1305
<i>Hongxia Yang and Quan Lu</i>	

Functional Regression with Mode-Sparsity Constraint	1311
<i>Pei Yang and Jingrui He</i>	
Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets	1317
<i>Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn Keogh</i>	
Deep Convolutional Factor Analyser for Multivariate Time Series Modeling	1323
<i>Chao Yuan and Amit Chakraborty</i>	
Incorporating Pre-Training in Long Short-Term Memory Networks for Tweets Classification	1329
<i>Shuhan Yuan, Xintao Wu, and Yang Xiang</i>	
Low-Rank Sparse Feature Selection for Patient Similarity Learning	1335
<i>Mengting Zhan, Shilei Cao, Buyue Qian, Shiyu Chang, and Jishang Wei</i>	
ROM: A Robust Online Multi-task Learning Approach	1341
<i>Chi Zhang, Peilin Zhao, Shuji Hao, Yeng Chai Soh, and Bu Sung Lee</i>	
Scalable Discrete Supervised Hash Learning with Asymmetric Matrix Factorization	1347
<i>Shifeng Zhang, Jianmin Li, Jinma Guo, and Bo Zhang</i>	
Multi-type Co-clustering of General Heterogeneous Information Networks via Nonnegative Matrix Tri-Factorization	1353
<i>Xianchao Zhang, Haixin Li, Wenxin Liang, and Jiebo Luo</i>	
Dynamic Poisson Factor Analysis	1359
<i>Yizhe Zhang, Yue Zhao, Lawrence David, Ricardo Henao, and Lawrence Carin</i>	
Gaussian Component Based Index for GMMs	1365
<i>Linfei Zhou, Bianca Wackersreuther, Frank Fiedler, Claudia Plant, and Christian Böhm</i>	
Multi-label Learning with Emerging New Labels	1371
<i>Yue Zhu, Kai-Ming Ting, and Zhi-Hua Zhou</i>	

Author Index