

2014 IEEE International Conference on Data Mining Workshop

(ICDMW 2014)

**Shenzhen, China
14 December 2014**

Pages 1-660



**IEEE Catalog Number: CFP1456B-POD
ISBN: 978-1-4799-4273-2**

2014 IEEE International Conference on Data Mining Workshop

ICDMW 2014

Table of Contents

Message from the Conference Chairs.....	xvii
Message from the Workshop Co-Chairs.....	xix
Preface to OEDM 2014.....	xxi
Preface to BioDM 2014.....	xxiii
Preface to DDDM 2014.....	xxvii
Preface to DMS 2014.....	xxix
Preface to SENTIRE 2014.....	xxxi
Preface to MoDAT 2014.....	xxxii
Preface to DaMNet 2014.....	xxxiv
Preface to HDM 2014.....	xxxvi
Preface to IClaNov 2014.....	xxxviii
Preface to DMBIH 2014.....	xli
Preface to BASNA 2014.....	xliii
Preface to SoftwareMining-2014.....	xlv
Preface to IDP 2014.....	xlvii
Preface to DMMI 2014.....	xlvix
Preface to SDA 2014.....	lii
Preface to DINA 2014.....	liv
Preface to SBDMV 2014.....	lvii

Ninth International Workshop on Spatial and Spatio-Temporal Data Mining—SSTD 2014

Spatio-Temporal Trajectory Region-of-Interest Mining Using Delaunay Triangulation	1
<i>Luke Bermingham, Kyungmi Lee, and Ickjai Lee</i>	
Identifying Team Style in Soccer Using Formations Learned from Spatiotemporal Tracking Data	9
<i>Alina Bialkowski, Patrick Lucey, Peter Carr, Yisong Yue, Sridha Sridharan, and Iain Matthews</i>	
Learning a Spatial Ensemble of Classifiers for Raster Classification: A Summary of Results	15
<i>Zhe Jiang, Shashi Shekhar, Azamat Kamzin, and Joseph Knight</i>	

Dynamic Time Warping Based on Cubic Spline Interpolation for Time Series Data Mining	19
<i>Hailin Li, Xiaoji Wan, Ye Liang, and Shile Gao</i>	
Why Checkins: Exploring User Motivation on Location Based Social Networks	27
<i>Fengjiao Wang, Guan Wang, and Philip S. Yu</i>	
Directional Higher Order Information for Spatio-Temporal Trajectory Dataset	35
<i>Ye Wang, Kyungmi Lee, and Ickjai Lee</i>	
Context-Aware Online Spatiotemporal Traffic Prediction	43
<i>Jie Xu, Dingxiong Deng, Ugur Demiryurek, Cyrus Shahabi, and Mihaela van der Schaar</i>	
Classification of 3D Surface Data Using the Concept of Vertex Unique Labelled Subgraphs	47
<i>Wen Yu, Frans Coenen, Michele Zito, and Kwankamon Dittakan</i>	

Ninth Workshop on Optimization Based Techniques for Emerging Data Mining—OEDM 2014

Measuring and Predicting Systemic Risk in the Chinese Banking System	55
<i>Yibing Chen, Yong Shi, Cheng-Few Lee, Minqiang Li, and Yuewen Liu</i>	
A Hierarchy Method Based on LDA and SVM for News Classification	60
<i>Limeng Cui, Fan Meng, Yong Shi, Minqiang Li, and An Liu</i>	
Deep Twin Support Vector Machine	65
<i>Dewei Li, Yingjie Tian, and Honggui Xu</i>	
Proximal Classifier via Absolute Value Inequalities	74
<i>Yuan-Hai Shao, Chun-Na Li, Zhen Wang, Ming-Zeng Liu, and Nai-Yang Deng</i>	
An Optimal Approach for Pruning Annular Regularized Extreme Learning Machines	80
<i>Lavneet Singh and Girija Chetty</i>	
Misclassification Minimization Based on Multiple Criteria Linear Programming	88
<i>Bo Wang, Yong Shi, Wayne Wei Huang, and Guanfeng Liu</i>	
Biased Support Vector Machine with Self-Constructed Universum for PU Learning	93
<i>Ying Zhang, Yingjie Tian, and Zhiquan Qi</i>	
Active Learning with Nonparallel Support Vector Machine for Binary Classification	101
<i>Xi Zhao, Zhensong Chen, and Yong Shi</i>	
Latent Factor SVM for Text Categorization	105
<i>Xiaofei Zhou, Li Guo, Ping Liu, and Yanbing Liu</i>	

Fifth Workshop on Biological Data Mining and Its Applications in Healthcare—BioDM 2014

Evaluation of the Relevance of Linear and Nonlinear Posturographic Features in the Recognition of Healthy Subjects and Stroke Patients	111
<i>Luiz Henrique Giovanini Marrega and Julio Cesar Nievola</i>	
Incorporating Spontaneous Reporting System Data to Aid Causal Inference in Longitudinal Healthcare Data	119
<i>Jenna M. Reps and Uwe Aickelin</i>	

Text Mining for Hypotheses and Results in Translational Medicine Studies	127
<i>Terry H. Tsai, Niels Kasch, Craig Pfeifer, and Tim Oates</i>	
Minimum Information Loss Based Multi-kernel Learning for Flagellar Protein Recognition in Trypanosoma Brucei	133
<i>Jim Jing-Yan Wang and Xin Gao</i>	

Seventh International Workshop on Domain Driven Data Mining—DDDM 2014

Wineinformatics: Applying Data Mining on Wine Sensory Reviews Processed by the Computational Wine Wheel	142
<i>Bernard Chen, Christopher Rhodes, Aaron Crawford, and Lorri Hambuchen</i>	
A Data Mining Framework to Model Consumer Indebtedness with Psychological Factors	150
<i>Jon Garibaldi, Eamonn Ferguson, Uwe Aickelin, and Jon Garibaldi</i>	
Analysis and Real-Time Prediction of Local Incident Impact on Transportation Networks	158
<i>Yuye He, Sebastien Blandin, Laura Wynter, and Barry Trager</i>	
NLMF: NonLinear Matrix Factorization Methods for Top-N Recommender Systems	167
<i>Santosh Kabbur and George Karypis</i>	
Two-Phase Attribute Ordering for Unsupervised Ranking of Multi-attribute Objects	175
<i>Chun-Guo Li, Xing Mei, and Bao-Gang Hu</i>	
Applying Cost-Sensitive Classification for Financial Fraud Detection under High Class-Imbalance	183
<i>Stephen O. Moepya, Sharat S. Akhoury, and Fulufhelo V. Nelwamondo</i>	
Moneyball for Academia: Toward Measuring and Maximizing Faculty Performance and Impact	193
<i>Andrew Nocka, Danning Zheng, Tianran Hu, and Jiebo Luo</i>	
A Standard Bibliography Recommended Method Based on Topic Model and Fusion of Multi-feature	198
<i>Fa Shao, Yan-Tuan Xian, Jian-Yi Guo, Zheng-Tao Yu, and Cun-Li Mao</i>	
Popular Items or Niche Items: Flexible Recommendation Using Cosine Patterns	205
<i>Yaqiong Wang, Junjie Wu, Zhiang Wu, Hua Yuan, and Xu Zhang</i>	
Clustering Ensemble and Application in HST Dataset	213
<i>Wenchao Xiao, Yan Yang, Hongjun Wang, and Yingge Xu</i>	
Mining Cursor Motions to Find the Gender, Experience, and Feelings of Computer Users	221
<i>Takashi Yamauchi and Casady Bowman</i>	
Statistical and Machine Learning Approach to Assessing the Environmental Impact on Walking Patterns	231
<i>Mingjing Yang, Huiru Zheng, Haiying Wang, Sally McClean, and Ruth E. Mayagoitia</i>	
Device Performance Analysis of Cloud Computing Data Center Two-Layer Interconnection Model Based on MPLS/IP Core Backbone Network	236
<i>Hanmin Ye, Zihang Song, and Qianting Sun</i>	
Discovering Organizational Correlations from Twitter	243
<i>Jingyuan Zhang, Xiaoxiao Shi, Xiangnan Kong, Hong-Han Shuai, and Philip S. Yu</i>	

Automatic Learning Common Definitional Patterns from Multi-domain Wikipedia Pages	251
<i>Jingsong Zhang, Yinglin Wang, and Dingyu Yang</i>	
Supervised Adaptive-Transfer PLSA for Cross-Domain Text Classification	259
<i>Rui Zhao and Kezhi Mao</i>	
Text Feature Selection Based on Class Subspace	267
<i>Xiaofei Zhou, Li Guo, Tianyi Wang, and Yue Hu</i>	

Fourth International Workshop on Data Mining for Service—DMS 2014

Understanding Changes in Customer Purchase Behavior: Study of Attenuation Model for Multiparametric Purchase Preferences	274
<i>Stephane Cheung, Yasuyuki Shirai, Hiroyuki Morita, and Masakazu Nakamoto</i>	
Visualization of Patient Distributions in a Hospital Based on the Clinical Actions Stored in EHR	284
<i>Shoji Hirano and Shusaku Tsumoto</i>	
Data Mining-Based Service Quality Estimation in Hospital Information System	289
<i>Toshihiko Kawamura, Tomohiro Kimura, and Shusaku Tsumoto</i>	
A Comparison of Algorithms for TV Ad Targeting	296
<i>Brendan Kitts and Dyng Au</i>	
Examination of Reliability of Missing Value Recovery in Data Mining	306
<i>Shigang Liu and Honghua Dai</i>	
Visual Exploration of a Series of Academic Conferences	314
<i>Kazuo Misue</i>	
Configurable Analytics Framework for Management of Enterprise IT Systems	321
<i>Maitreya Natu, Vaishali Sadaphal, and Vaishali Kulkarni</i>	
Factors of Patient Satisfaction Based on Distant Analysis in HCAHPS Databases	328
<i>Masumi Okuda, Akira Yasuda, and Shusaku Tsumoto</i>	
Personalising Mobile Advertising Based on Users' Installed Apps	338
<i>Jenna Reys, Uwe Aickelin, Jonathan Garibaldi, and Chris Damski</i>	
An Incremental Tensor Factorization Approach for Web Service Recommendation	346
<i>Wancai Zhang, Hailong Sun, Xudong Liu, Xiaohui, and Guo</i>	

Sentiment Elicitation from Natural Text for Information Retrieval and Extraction—SENTIRE

A Hybrid Approach for Emotion Detection in Support of Affective Interaction	352
<i>Sonja Gievska, Kiril Koroveshovski, and Tatjana Chavdarova</i>	
Commonsense Knowledge as the Glue in a Hybrid Model of Computational Creativity	360
<i>Andrés Gómez de Silva Garza, Erik Cambria, and Rafael Pérez Y. Pérez</i>	
Social Preference Ontologies for Enriching User and Item Data in Recommendation Systems	365
<i>Christopher Krauss and Stefan Arbanowski</i>	
Cyberbullying Detection using Time Series Modeling	373
<i>Nektaria Potha and Manolis Maragoudakis</i>	

Emotion Recognition from Text Based on Automatically Generated Rules	383
<i>Shadi Shaheen, Wassim El-Hajj, Hazem Hajj, and Shady Elbassuoni</i>	
Semi-Supervised Method for Multi-category Emotion Recognition in Tweets	393
<i>Valentina Sintsova, Claudiu Musat, and Pearl Pu</i>	
A Localization Toolkit for Sentic Net	403
<i>Yunqing Xia, Xiaoyu Li, Erik Cambria, and Amir Hussain</i>	
Towards Summarizing Popular Information from Massive Tourism Blogs	409
<i>Hua Yuan, Hualin Xu, Yu Qian, and Kai Ye</i>	
Joint Propagation and Refinement for Mining Opinion Words and Targets	417
<i>Qiyun Zhao, Hao Wang, and Pin Lv</i>	
A Semi-Supervised Self-Adaptive Classifier over Opinionated Streams	425
<i>Max Zimmerann, Eirini Ntoutsi, and Myra Spiliopoulou</i>	
 Designing the Market of Data for Practical Data Sharing via Educational and Innovative Communications—MoDAT 2014	
Abduction Dealing with Potential Values and Its Datasets	433
<i>Akinori Abe</i>	
Estimation of Novelty Assessment of Strategic Scenarios Using Relativeness	441
<i>Teruaki Hayashi and Yukio Ohsawa</i>	
Modeling of Writing and Thinking Process in Handwriting by Digital Pen Analysis	447
<i>Kenshin Ikegami and Yukio Ohsawa</i>	
A Narrative Analysis by Text Mining Technique Using Key Graph: Similarity and Difference of a View of Oral Health and Oral Risk Cognition between Japanese Living People and Dentists	455
<i>Fukiko Kobayashi and Yumiko Nara</i>	
A System Design Method Based on Interaction between Logic and Data Sets	462
<i>Noriyuki Kushiro, Shodai Matsuda, Ryoichi Torikai, and Kunio Takahara</i>	
Creating Essential Hypothesis and Rules in Product Planning—Introducing Abductive Inference Model	470
<i>Jun Nakamura</i>	
Exploring Markets of Data for Personal Health Information	477
<i>K. Thomas Pickard</i>	
Domain-Independent Unsupervised Text Segmentation for Data Management	481
<i>Makoto Sakahara, Shogo Okada, and Katsumi Nitta</i>	
Mining Interesting Meta-Paths from Complex Heterogeneous Information Networks	488
<i>Baoxu Shi and Tim Weninger</i>	
Adaptability of Dynamic Quality Function Deployment to the Concept Design of Electric Trucks: An Attempt to Knowledge Management and Data Engineering	496
<i>Masahiko Teramoto and Jun Nakamura</i>	
A Parallel Implementation of Idea Graph to Extract Rare Chances from Big Data	503
<i>Qinyong Wang, Hao Wang, Chen Zhang, Wei Wang, Zhe Chen, and Fanjiang Xu</i>	

Introduction of Search Engine Focusing on Trend-Related Queries to Market of Data	511
<i>YanJun Zhu, Yasufumi Takama, Yu Kato, Shogo Kori, and Hiroshi Ishikawa</i>	

Fourth International Workshop on Data Mining in Networks—DaMNet 2014

Quick Mining of Isomorphic Exact Large Patterns from Large Graphs	517
<i>Islam Almasri, Xin Gao, and Nina Fedoroff</i>	
Online Fusion of Incremental Learning for Wireless Sensor Networks	525
<i>H.H.W.J. Bosman, G. Iacca, H.J. Wörtche, and A. Liotta</i>	
EgoLP: Fast and Distributed Community Detection in Billion-Node Social Networks	533
<i>Nazar Buzun, Anton Korshunov, Valeriy Avanesov, Ilya Filonenko, Ilya Kozlov, Denis Turdakov, and Hangkyu Kim</i>	
Recovering Human Mobility Flow Models and Daily Routine Patterns in a Smart Environment	541
<i>Li Chen and William K. Cheung</i>	
A Link Prediction Method That Can Learn from Network Dynamics	549
<i>Yang Chen, Ke-Jia Chen, and Yun Li</i>	
Threatening Privacy across Social Graphs: A Structural Features Approach	554
<i>Priya Govindan, Tina Eliassi-Rad, Jin Xu, Shawndra Hill, and Chris Volinsky</i>	
Identify Influential Social Network Spreaders	562
<i>Chung-Yuan Huang, Yu-Hsiang Fu, and Chuen-Tsai Sun</i>	
R-HBase: A Multi-dimensional Indexing Framework for Cloud Computing Environment	569
<i>Shan Huang, Botao Wang, Jingyang Zhu, Guoren Wang, and Ge Yu</i>	
Defensibility-Based Classification for Argument Mining	575
<i>Hiroyuki Kido and Yukio Ohsawa</i>	
Privacy-Preserving Queries over Outsourced Data with Access Pattern Protection	581
<i>Shun-Pun Li and Man-Hon Wong</i>	
Community Detection on Large Graph Datasets for Recommender Systems	589
<i>Rohit Parimi and Doina Caragea</i>	
Expander Graph Quality Optimisation in Randomised Communication	597
<i>Pasu Poonpakdee and Giuseppe Di Fatta</i>	
Node-Specific Triad Pattern Mining for Complex-Network Analysis	605
<i>Marco Winkler and Joerg Reichardt</i>	
Exploiting Paper Contents and Citation Links to Identify and Characterise Specialisations	613
<i>Han Xu, Eric Martin, and Ashesh Mahidadia</i>	

Second International Workshop on High Dimensional Data Mining Duration—HDM 2014

SUBSCALE: Fast and Scalable Subspace Clustering for High Dimensional Data	621
<i>Amardeep Kaur and Amitava Datta</i>	
Random KNN	629
<i>Shengqiao Li, E. James Harner, and Donald A. Adjeroh</i>	

Out-of-Sample Error Estimation: The Blessing of High Dimensionality	637
<i>Luca Oneto, Alessandro Ghio, Sandro Ridella, Jorge Luis Reyes Ortiz, and Davide Anguita</i>	
Who Wrote This? Textual Modeling with Authorship Attribution in Big Data	645
<i>Naruemon Pratanwanich and Pietro Lio</i>	
Two approaches of using heavy tails in high dimensional EDA	653
<i>Momodou L. Sanyang, Hanno Muehlbrandt, and Ata Kaban</i>	
High Dimensional Matrix Relevance Learning	661
<i>Frank-Michael Schleich, Thomas Villmann, and Xibin Zhu</i>	
Dimensionality Reduction Based Similarity Visualization for Neural Gas	668
<i>Kadim Taşdemir</i>	
Boosting for Vote Learning in High-Dimensional kNN Classification	676
<i>Nenad Tomašev</i>	
Adaptive Semi-Supervised Dimensionality Reduction	684
<i>Jia Wei, Jiabing Wang, Qianli Ma, and Xuan Wang</i>	
Rényi Divergence Based Generalization for Learning of Classification Restricted Boltzmann Machines	692
<i>Qian Yu, Yuexian Hou, Xiaozhao Zhao, and Guochen Cheng</i>	

Incremental Classification Concept Drift and Novelty Detection—IClaNov 2014

Efficient Anomaly Detection by Isolation Using Nearest Neighbour Ensemble	698
<i>Tharindu R. Bandaragoda, Kai Ming Ting, David Albrecht, Fei Tony Liu, and Jonathan R. Wells</i>	
Merging Classifiers of Different Classification Approaches	706
<i>Antonina Danylenko and Welf Löwe</i>	
Incremental Ensemble Classifier Addressing Non-stationary Fast Data Streams	716
<i>Brandon S. Parker, Latifur Khan, and Albert Bifet</i>	
Drift Detection for Multi-label Data Streams Based on Label Grouping and Entropy	724
<i>Zhongwei Shi, Yimin Wen, Chao Feng, and Hai Zhao</i>	

Data Mining in Biomedical Informatics and Healthcare—DMBIH 2014

Mining Top-K Frequent Closed Patterns from Gene Expression Data	732
<i>Shufan Ji, Xuejiao Wang, Yi Zong, and Xiaopeng Gao</i>	
Rough-Set-Based ADR Signaling from Spontaneous Reporting Data with Missing Values	740
<i>Wen-Yang Lin, Lin Lan, and Fong-Sheng Huang</i>	
Glucoguide: An Intelligent Type-2 Diabetes Solution Using Data Mining and Mobile Computing	748
<i>Yan Luo, Charles Ling, Jody Schuurman, and Robert Petrella</i>	
Imputation of Missing Values in Time Series with Lagged Correlations	753
<i>Shah Atiqur Rahman, Yuxiao Huang, Jan Claassen, and Samantha Kleinberg</i>	

Refining Adverse Drug Reactions Using Association Rule Mining for Electronic Healthcare Data	763
<i>Jenna M. Reps, Uwe Aickelin, Jiangang Ma, and Yanchun Zhang</i>	
Towards Achieving Diagnostic Consensus in Medical Image Interpretation	771
<i>Mike Seidel, Alexander Rasin, Jacob D. Furst, and Daniela S. Raicu</i>	

Fifth International Workshop on Business Applications of Social Network Analysis—BASNA 2014

An Improved Collaborative Method for Recommendation and Rating Prediction	781
<i>Guoyong Cai, Rui Lv, Hao Wu, and Xia Hu</i>	
A Business Intelligence Solution for Frequent Pattern Mining on Social Networks	789
<i>Fan Jiang and Carson Kai-Sang Leung</i>	
NaviSoc: A Socially Enhanced Real-Time Navigator	797
<i>Nikolaos Louloudakis, Vasileios Theodosiadis, Haridimos Kondylakis, and Kostas Stefanidis</i>	
Multiple Subscriber-Identity-Module Detection Using Social Network Analysis Techniques	804
<i>Ahmed Maher and Ahmed Ghoneim</i>	
Email Analytics for Support Center Performance Analysis	810
<i>Kunal Ranjan and Lipika Dey</i>	
A Scalable Algorithm for Discovering Topologies in Social Networks	818
<i>Jyoti Rani Yadav, D.V.L.N. Somayajulu, and P. Radha Krishna</i>	

Third International Workshop on Software Mining—SoftwareMining-2014

Software Process Simulation Based on Mining Software Repositories	828
<i>Verena Honsel, Daniel Honsel, and Jens Grabowski</i>	
Ontology-Based Software Requirements Reuse: Case Study in Fire Control Software Product Line Domain	832
<i>Elif Kamer Karataş, Barış İyidir, and Ayşenur Birtürk</i>	
Taming a Fuzzer Using Delta Debugging Trails	840
<i>Yuanli Pei, Arpit Christi, Xiaoli Fern, Alex Groce, and Weng-Keen Wong</i>	
A Comparison of Approaches to Chinese Word Segmentation in Hadoop	844
<i>Zhangang Wang and Bangjie Meng</i>	
What Makes an Open Source Code Popular on Git Hub?	851
<i>Simon Weber and Jiebo Luo</i>	

Third International Workshop on Intelligent Data Processing Online Social Networks and E-Health Data

Medical Error Prevention Based on Path Integration System Approach	856
<i>Sheung Wai Chan, Clement Leung, Victor Cheng, and Jiming Liu</i>	
Illumination Invariant Face Recognition with Particle Swarm Optimization	862
<i>Yu Cheng, Zhigang Jin, Cunming Hao, and Xingsen Li</i>	
An Effective Clustering Algorithm for Auto-Detecting Well-Separated Clusters	867
<i>Jinyuan He, Gansen Zhao, Hao Lan Zhang, Kotagiri Ramamohanarao, and Chaoyi Pang</i>	

An Extenics-Based Criteria Clustering Method	875
<i>Xingsen Li, Haolan Zhang, and Wei Deng</i>	
Accelerated Online Learning for Collaborative Filtering and Recommender Systems	879
<i>Yuan-Xiang Li, Zhi-Jie Li, Feng Wang, and Li Kuang</i>	
Query Hidden Attributes in Social Networks	886
<i>Azade Nazi, Saravanan Thirumuruganathan, Vagelis Hristidis, Nan Zhang, Khaled Shaban, and Gautam Das</i>	
Exploring Mixed Membership Stochastic Block Models via Non-negative Matrix Factorization	892
<i>Chengbin Peng and Ka-Chun Wong</i>	
A Collaborative and Adaptive Intrusion Detection Based on SVMs and Decision Trees	898
<i>Luyao Teng, Shaohua Teng, Feiyi Tang, Haibin Zhu, Wei Zhang, Dongning Liu, and Lu Liang</i>	
An Efficient Algorithm for Finding Continuous Coherent Evolution Bicluster in Time-Series Data	906
<i>Yun Xue, Jie Luo, Haolan Zhang, Zhengling Liao, Meihang Li, and Qiuhua Kuang</i>	

First COOL-SNA Workshop on Connecting Online and Offline Social Network Analysis

Categorizing Social Multimedia by Neighborhood Decision Using Local Pairwise Label Correlation	913
<i>Jun Huang, Guorong Li, Shuhui Wang, and Qingming Huang</i>	
Clustering and Classification of Like-Minded People from their Tweets	921
<i>Soufiene Jaffali, Salma Jamoussi, Abdelmajid Ben Hamadou, and Kamel Smaili</i>	
Chinese Microblog Sentiment Classification Based on Deep Belief Nets with Extended Multi-Modality Features	928
<i>Xiao Sun, Chengcheng Li, Wanyi Xu, and Fuji Ren</i>	
Estimating Online User Location Distribution without GPS Location	936
<i>Yusheng Xie, Yu Cheng, Ankit Agrawal, and Alok Choudhary</i>	
Cost-Effective User Monitoring for Popularity Prediction of Online User-Generated Content	944
<i>Mengmeng Yang, Kai Chen, Zhongchen Miao, and Xiaokang Yang</i>	
A Retweet Number Prediction Model Based on Followers' Retweet Intention and Influence	952
<i>Huidong Zhao, Gang Liu, Chuan Shi, and Bin Wu</i>	

The IEEE ICDM Workshop on Data Mining in Medical Imaging—DMMI 2014

Blood Vessel Segmentation in Pathological Retinal Image	960
<i>Zhe Han, Yilong Yin, Xianjing Meng, Gongping Yang, and Xiaowei Yan</i>	
Evaluating Spectral Unmixing Quality in the Absence of Reference Objects for Time Lapse Fluorescence Microscopy	968
<i>Liyan Liu, Andrey Kan, Jie H.S. Zhou, John F. Markham, Philip D. Hodgkin, and Christopher Leckie</i>	
Segmentation and Splitting of Touching Vaginal Bacteria Based on Superpixel and Effective Distance	976
<i>Youyi Song, Dong Ni, Liang He, Siping Chen, Baiying Lei, and Tianfu Wang</i>	

Mining of Training Samples for Multiple Learning Machines in Computer-Aided Detection of Lesions in CT Images	982
<i>Kenji Suzuki</i>	
A Novel Time-Frequency Analysis Approach for Nonstationary Time Series Using Multiresolution Wavelet	990
<i>Si-Rui Tan, Yang Li, and Ke Li</i>	
Combining Multiple Network Features for Mild Cognitive Impairment Classification	996
<i>Lipeng Wang, Fei Fei, Biao Jie, and Daoqiang Zhang</i>	

Social Multimedia Data Mining—SMMDM 2014

Recommender Systems Using Harmonic Analysis	1004
<i>Gilbert Badaro, Hazem Hajj, Ali Haddad, Wassim El-Hajj, and Khaled Bashir Shaban</i>	
Topical Influential User Analysis with Relationship Strength Estimation in Twitter	1012
<i>Xinyue Liu, Hua Shen, Fenglong Ma, and Wenxin Liang</i>	
Semantic Features for Food Image Recognition with Geo-Constraints	1020
<i>Xinhang Song, Shuqiang Jiang, Ruihan Xu, and Luis Herranz</i>	
The Eyes of the Beholder: Gender Prediction Using Images Posted in Online Social Networks	1026
<i>Quanzeng You, Sumit Bhatia, Tong Sun, and Jiebo Luo</i>	

Workshop on Scalable Data Analytics: Theory and Applications—SDA 2014

A Further Investigation on the Reliability of Extreme Learning Machines	1031
<i>Yanxing Hu, Yuan Wang, Jane Jia You, Jame N.K. Liu, and Yulin He</i>	
DISC: Efficient Uncertain Frequent Pattern Mining with Tightened Upper Bounds	1038
<i>Richard Kyle MacKinnon, Teagan D. Strauss, and Carson Kai-Sang Leung</i>	
Parallel Frequent Pattern Mining without Candidate Generation on GPUs	1046
<i>Fei Wang and Bo Yuan</i>	
A New Fast Minimum Spanning Tree-Based Clustering Technique	1053
<i>Xiaochun Wang, Xia L. Wang, and Jihua Zhu</i>	
HDminer: Efficient Mining of High Dimensional Frequent Closed Patterns from Dense Data	1061
<i>Jianpeng Xu and Shufan Ji</i>	

Data Integration and Applications—DINA 2014

Analytics-Enabled Disaster Preparedness and Emergency Management	1068
<i>Karen España, Rhia Trogo, James Faeldon, Juanito Santiago, and Delfin Jay Sabido</i>	
Integration and Automation of Data Preparation and Data Mining	1076
<i>Shrikanth Narayanan, Ayush Jaiswal, Yao-Yi Chiang, Yanhui Geng, Craig A. Knoblock, and Pedro Szekely</i>	
Comparing Classifiers in Historical Census Linkage	1086
<i>Laura Richards, Luiza Antonie, Shawki Areibi, Gary Grewal, Kris Inwood, and J. Andrew Ross</i>	
Semi-Supervised Consensus Clustering: Reducing Human Effort	1095
<i>Tobias Vogel and Felix Naumann</i>	

Iteratively Learning Conditional Statements in Transforming Data by Example	1105
<i>Bo Wu and Craig A. Knoblock</i>	

Spatial Big Data Mining and Visualization—SBDMV 2014

Video Retrieval Methods Using Geographic Information in Windows Azure Cloud	1113
<i>Zhigang Han, Yunfeng Kong, Qin Fen, and Pinde Fu</i>	
A Novel Dummy-Based Mechanism to Protect Privacy on Trajectories	1120
<i>Xichen Wu and Guangzhong Sun</i>	
Estimating the Distribution of Economy Activity: A Case Study in Jiangsu Province (China) Using Large Scale Social Network Data	1126
<i>Chang Xiaomeng, Lu Guozhen, Yue Yang, and Li Qingquan</i>	
A Novel Disparity Refinement Method Based on Semi-Global Matching Algorithm	1135
<i>Yuechao Xie, Siyu Zeng, and Long Chen</i>	
Real-time Dynamic Visualization Techniques for Massive Geospatial Data	1143
<i>Zhou Yanan, Hu Xiaodong, and Li Jun</i>	
Discretizing Numerical Attributes in Decision Tree for Big Data Analysis	1150
<i>Yiqun Zhang and Yiu-Ming Cheung</i>	

PhD Forum

Leverage Item Popularity and Recommendation Quality via Cost-Sensitive Factorization Machines	1158
<i>Chih-Ming Chen, Hsin-Ping Chen, Ming-Feng Tsai, and Yi-Hsuan Yang</i>	
Dual Uncertainty Minimization Regularization and Its Applications on Heterogeneous Data	1163
<i>Yu Cheng, Alok Choudhary, Jun Wang, Sharath Pankanti, and Huan Liu</i>	
Efficient and Accelerated Online Learning for Sparse Group Lasso	1171
<i>Li Zhi-Jie, Li Yuan-Xiang, Wang Feng, Yu Fei, and Xiang Zheng-Long</i>	
Sequential Pattern Analysis with Right Granularity	1178
<i>Chuanren Liu, Kai Zhang, and Hui Xiong</i>	
Clinical Decision Making: A Framework for Predicting Rx Response	1185
<i>Aarti Sathyanarayana, Jyotishman Pathak, Rozalina McCoy, Santiago Romero-Brufau, Maryam Panaziah, and Jaideep Srivastava</i>	
Joint Visual and Textual Mining on Social Media	1189
<i>Jia Xu</i>	
Fuzzy Sentiment Membership Determining for Sentiment Classification	1191
<i>Chuanjun Zhao, Suge Wang, and Deyu Li</i>	

Demo Papers

Big Data Stream Learning with SAMOA	1199
<i>Albert Bifet and Gianmarco De Francisci Morales</i>	
RQL: A SQL-Like Query Language for Discovering Meaningful Rules	1203
<i>Brice Chardin, Emmanuel Coquery, Marie Pailloux, and Jean-Marc Petit</i>	

Trip Router: A Time-Sensitive Route Recommender System	1207
<i>Hsun-Ping Hsieh, Cheng-Te Li, and Shou-De Lin</i>	
Eventera: Real-Time Event Recommendation System from Massive Heterogeneous Online Media	1211
<i>Dongyeop Kang, Donggyun Han, Nahea Park, Sangtae Kim, U. Kang, and Soobin Lee</i>	
Finding Valuable Yelp Comments by Personality, Content, Geo, and Anomaly Analysis	1215
<i>Jay Koven, Hossein Siadati, and Ching-Yung Lin</i>	
Pathway-Finder: An Interactive Recommender System for Supporting Personalized Care Pathways	1219
<i>Rui Liu, Raj Velamur Srinivasan, Kiyana Zolfaghar, Si-Chi Chin, Senjuti Basu Roy, Aftab Hasan, and David Hazel</i>	
Visual-VM: A Social Network Visualization Tool for Viral Marketing	1223
<i>Cheng Long and Raymond Chi-Wing Wong</i>	
HealthSCOPE: An Interactive Distributed Data Mining Framework for Scalable Prediction of Healthcare Costs	1227
<i>Ames Marquardt, Stacey Newman, Deepa Hattarki, Rajagopalan Srinivasan, Shanu Sushmita, Prabhu Ram, Viren Prasad, David Hazel, Archana Ramesh, Martine De Cock, and Ankur Teredesai</i>	
e2-Diagnoser: A System for Monitoring, Forecasting and Diagnosing Energy Usage	1231
<i>Joern Ploennigs, Bei Chen, Paulito Palmes, and Raymond Lloyd</i>	
Smart Saver: A Consumer-Oriented Web Service for Energy Disaggregation	1235
<i>Guoming Tang, Jie Chen, Cheng Chen, and Kui Wu</i>	
Web Service QoS Prediction Approach in Mobile Internet Environments	1239
<i>Lubao Wang, Qibo Sun, Shangguang Wang, You Ma, Jinliang Xu, and Jinglin Li</i>	
Interoperability-Enriched App Recommendation	1242
<i>Shi Wenxuan and Yin Airu</i>	
RepComment: A Fair Comment-Sentiment Representation System	1246
<i>Ting Wu, Chun Yin Tan, Ming Yip Cheung, and Pan Hui</i>	
LibEDM: A Platform for Ensemble Based Data Mining	1250
<i>Qiangli Zhao and Yanhuang Jiang</i>	
A New Coupled Metric Learning for Real-time Anomalies Detection with High-Frequency Field Programmable Gate Arrays	1254
<i>Frank Jiang and Dan Luo</i>	

Author Index