Contents – Part II

Spatiotemporal and Image Data

Denoising Time Series by Way of a Flexible Model for Phase Space
Reconstruction ................................................................. 3
   Minhazul Islam Sk and Arunava Banerjee

Distributed Sequential Pattern Mining in Large Scale Uncertain Databases . . . . . . . 17
   Jiaqi Ge and Yuni Xia

DeepCare: A Deep Dynamic Memory Model for Predictive Medicine ............. 30
   Trang Pham, Truyen Tran, Dinh Phung, and Svetha Venkatesh

Indoor Positioning System for Smart Homes Based on Decision Trees
and Passive RFID ............................................................ 42
   Frédéric Bergeron, Kevin Bouchard, Sébastien Gaboury,
   Sylvain Giroux, and Bruno Bouchard

Deep Feature Extraction from Trajectories for Transportation
Mode Estimation ............................................................. 54
   Yuki Endo, Hiroyuki Toda, Kyosuke Nishida, and Akihisa Kawanobe

Online Learning for Accurate Real-Time Map Matching ......................... 67
   Biwei Liang, Tengjiao Wang, Shun Li, Wei Chen, Hongyan Li,
   and Kai Lei

Multi-hypergraph Incidence Consistent Sparse Coding for Image
Data Clustering .............................................................. 79
   Xiaodong Feng, Sen Wu, Wenjun Zhou, and Zhiwei Tang

Robust Multi-view Manifold Ranking for Image Retrieval ....................... 92
   Jun Wu, Jianbo Yuan, and Jiebo Luo

Image Representation Optimization Based on Locally Aggregated
Descriptors ................................................................. 104
   Shijiang Chen, Guiguang Ding, Chenxiao Li, and Yuchen Guo

Reusing Extracted Knowledge in Genetic Programming to Solve Complex
Texture Image Classification Problems .................................. 117
   Muhammad Iqbal, Bing Xue, and Mengjie Zhang
Personal Credit Profiling via Latent User Behavior Dimensions on Social Media ........................................... 130
Guangming Guo, Feida Zhu, Enhong Chen, Le Wu, Qi Liu, Yingling Liu, and Minghui Qiu

Linear Upper Confidence Bound Algorithm for Contextual Bandit Problem with Piled Rewards ........................................... 143
Kuan-Hao Huang and Hsuan-Tien Lin

Incremental Hierarchical Clustering of Stochastic Pattern-Based Symbolic Data ........................................... 156
Xin Xu, Jiaheng Lu, and Wei Wang

Computing Hierarchical Summary of the Data Streams ........................................... 168
Zubair Shah, Abdun Naser Mahmood, and Michael Barlow

Anomaly Detection and Clustering

Unsupervised Parameter Estimation for One-Class Support Vector Machines ........................................... 183
Zahra Ghafoori, Sutharshan Rajasegarar, Sarah M. Erfani, Shanika Karunasekera, and Christopher A. Leckie

Frequent Pattern Outlier Detection Without Exhaustive Mining ........................................... 196
Arnaud Giacometti and Arnaud Soulet

Ensembles of Interesting Subgroups for Discovering High Potential Employees ........................................... 208
Girish Keshav Palshikar, Kuleshwar Sahu, and Rajiv Srivastava

Dynamic Grouped Mixture Models for Intermittent Multivariate Sensor Data ........................................... 221
Naoya Takeishi, Takehisa Yairi, Naoki Nishimura, Yuta Nakajima, and Noboru Takata

Parallel Discord Discovery ........................................... 233
Tian Huang, Yongxin Zhu, Yishu Mao, Xinyang Li, Mengyun Liu, Yafei Wu, Yajun Ha, and Gillian Dobbie

Dboost: A Fast Algorithm for DBSCAN-Based Clustering on High Dimensional Data ........................................... 245
Yuxiao Zhang, Xiaorong Wang, Bingyang Li, Wei Chen, Tengjiao Wang, and Kai Lei

A Precise and Robust Clustering Approach Using Homophilic Degrees of Graph Kernel ........................................... 257
Haolin Yang, Deli Zhao, Lele Cao, and Fuchun Sun
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constraint Based Subspace Clustering for High Dimensional Uncertain Data</td>
<td>271</td>
</tr>
<tr>
<td>Xianchao Zhang, Lu Gao, and Hong Yu</td>
<td></td>
</tr>
<tr>
<td>A Clustering-Based Framework for Incrementally Repairing Entity Resolution</td>
<td>283</td>
</tr>
<tr>
<td>Qing Wang, Jingyi Gao, and Peter Christen</td>
<td></td>
</tr>
<tr>
<td>Adaptive Seeding for Gaussian Mixture Models</td>
<td>296</td>
</tr>
<tr>
<td>Johannes Blömer and Kathrin Bujna</td>
<td></td>
</tr>
<tr>
<td>A Greedy Algorithm to Construct L1 Graph with Ranked Dictionary</td>
<td>309</td>
</tr>
<tr>
<td>Shuchu Han and Hong Qin</td>
<td></td>
</tr>
</tbody>
</table>

**Novel Models and Algorithms**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Rule Based Open Information Extraction Method Using Cascaded Finite-State Transducer</td>
<td>325</td>
</tr>
<tr>
<td>Hailun Lin, Yuanzhuo Wang, Peng Zhang, Weiping Wang, Yinliang Yue, and Zheng Lin</td>
<td></td>
</tr>
<tr>
<td>Active Learning Based Entity Resolution Using Markov Logic</td>
<td>338</td>
</tr>
<tr>
<td>Jeffrey Fisher, Peter Christen, and Qing Wang</td>
<td></td>
</tr>
<tr>
<td>Modeling Adversarial Learning as Nested Stackelberg Games.</td>
<td>350</td>
</tr>
<tr>
<td>Yan Zhou and Murat Kantarcioglu</td>
<td></td>
</tr>
<tr>
<td>Fast and Semantic Measurements on Collaborative Tagging Quality</td>
<td>363</td>
</tr>
<tr>
<td>Yuqing Sun, Haiqi Sun, and Reynold Cheng</td>
<td></td>
</tr>
<tr>
<td>Matrices, Compression, Learning Curves: Formulation, and the GROUPNTEACH Algorithms</td>
<td>376</td>
</tr>
<tr>
<td>Bryan Hooi, Hyun Ah Song, Evangelos Papalexakis, Rakesh Agrawal, and Christos Faloutsos</td>
<td></td>
</tr>
<tr>
<td>Privacy Aware K-Means Clustering with High Utility</td>
<td>388</td>
</tr>
<tr>
<td>Thanh Dai Nguyen, Sunil Gupta, Santu Rana, and Svetla Venkatesh</td>
<td></td>
</tr>
<tr>
<td>Secure k-NN Query on Encrypted Cloud Data with Limited Key-Disclosure and Offline Data Owner</td>
<td>401</td>
</tr>
<tr>
<td>Youwen Zhu, Zhikuan Wang, and Yue Zhang</td>
<td></td>
</tr>
<tr>
<td>Hashing-Based Distributed Multi-party Blocking for Privacy-Preserving Record Linkage</td>
<td>415</td>
</tr>
<tr>
<td>Thilina Ranbaduge, Dinusha Vatsalan, Peter Christen, and Vassilios Verykios</td>
<td></td>
</tr>
</tbody>
</table>
### Text Mining and Recommender Systems

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enabling Hierarchical Dirichlet Processes to Work Better for Short Texts at Large Scale</td>
<td>431</td>
</tr>
<tr>
<td><strong>Khai Mai, Sang Mai, Anh Nguyen, Ngo Van Linh, and Khoat Than</strong></td>
<td></td>
</tr>
<tr>
<td>Query-Focused Multi-document Summarization Based on Concept Importance</td>
<td>443</td>
</tr>
<tr>
<td><strong>Hai-Tao Zheng, Ji-Min Guo, Yong Jiang, and Shu-Tao Xia</strong></td>
<td></td>
</tr>
<tr>
<td>Mirror on the Wall: Finding Similar Questions with Deep Structured Topic Modeling</td>
<td>454</td>
</tr>
<tr>
<td><strong>Arpita Das, Manish Shrivastava, and Manoj Chinnakotla</strong></td>
<td></td>
</tr>
<tr>
<td>An Efficient Dynamic Programming Algorithm for STR-IC-STR-IC-LCS Problem</td>
<td>466</td>
</tr>
<tr>
<td><strong>Daxin Zhu, Yingjie Wu, and Xiaodong Wang</strong></td>
<td></td>
</tr>
<tr>
<td>Efficient Page-Level Data Extraction via Schema Induction and Verification</td>
<td>478</td>
</tr>
<tr>
<td><strong>Chia-Hui Chang, Tian-Sheng Chen, Ming-Chuan Chen, and Jhung-Li Ding</strong></td>
<td></td>
</tr>
<tr>
<td>Transfer-Learning Based Model for Reciprocal Recommendation</td>
<td>491</td>
</tr>
<tr>
<td><strong>Chia-Hsin Ting, Hung-Yi Lo, and Shou-De Lin</strong></td>
<td></td>
</tr>
<tr>
<td>Enhanced SVD for Collaborative Filtering</td>
<td>503</td>
</tr>
<tr>
<td><strong>Xin Guan, Chang-Tsun Li, and Yu Guan</strong></td>
<td></td>
</tr>
<tr>
<td>Social Group Based Video Recommendation Addressing the Cold-Start Problem</td>
<td>515</td>
</tr>
<tr>
<td><strong>Chunfeng Yang, Yipeng Zhou, Liang Chen, Xiaopeng Zhang, and Dah Ming Chiu</strong></td>
<td></td>
</tr>
<tr>
<td>FeRoSA: A Faceted Recommendation System for Scientific Articles</td>
<td>528</td>
</tr>
<tr>
<td><strong>Tanmoy Chakraborty, Amrith Krishna, Mayank Singh, Niloy Ganguly, Pawan Goyal, and Animesh Mukherjee</strong></td>
<td></td>
</tr>
<tr>
<td>Dual Similarity Regularization for Recommendation</td>
<td>542</td>
</tr>
<tr>
<td><strong>Jing Zheng, Jian Liu, Chuan Shi, Fu Zen Zhuang, Jingzhi Li, and Bin Wu</strong></td>
<td></td>
</tr>
<tr>
<td>Collaborative Deep Ranking: A Hybrid Pair-Wise Recommendation Algorithm with Implicit Feedback</td>
<td>555</td>
</tr>
<tr>
<td><strong>Haochao Ying, Liang Chen, Yuwen Xiong, and Jian Wu</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Author Index**                                                                 | 569  |
Contents – Part I

Classification

Joint Classification with Heterogeneous Labels Using Random Walk with Dynamic Label Propagation ............................................. 3
  Yongxin Liao, Shenxi Yuan, Jian Chen, Qingyao Wu, and Bin Li

Hybrid Sampling with Bagging for Class Imbalance Learning ............ 14
  Yang Lu, Yiu-ming Cheung, and Yuan Yan Tang

Sparse Adaptive Multi-hyperplane Machine .................................. 27
  Khanh Nguyen, Trung Le, Vu Nguyen, and Dinh Phung

Exploring Heterogeneous Product Networks for Discovering Collective Marketing Hyping Behavior .......................................... 40
  Qinzhe Zhang, Qin Zhang, Guodong Long, Peng Zhang, and Chengqi Zhang

Optimal Training and Efficient Model Selection for Parameterized Large Margin Learning ......................................................... 52
  Yuxun Zhou, Jae Yeon Baek, Dan Li, and Costas J. Spanos

Locally Weighted Ensemble Learning for Regression ........................ 65
  Man Yu, Zongxia Xie, Hong Shi, and Qinghua Hu

Reliable Confidence Predictions Using Conformal Prediction ............ 77
  Henrik Linusson, Ulf Johansson, Henrik Boström, and Tuve Löfström

Grade Prediction with Course and Student Specific Models ............... 89
  Agoritsa Polyzou and George Karypis

Flexible Transfer Learning Framework for Bayesian Optimisation ........ 102
  Tinu Theckel Joy, Santu Rana, Sunil Kumar Gupta, and Svetla Venkatesh

A Simple Unlearning Framework for Online Learning Under Concept Drifts ................................................................. 115
  Sheng-Chi You and Hsuan-Tien Lin

User-Guided Large Attributed Graph Clustering with Multiple Sparse Annotations ............................................................... 127
  Jianping Cao, Senzhang Wang, Fengcai Qiao, Hui Wang, Feiyue Wang, and Philip S. Yu
Early-Stage Event Prediction for Longitudinal Data......................... 139  
Mahtab J. Fard, Sanjay Chawla, and Chandan K. Reddy

Toxicity Prediction in Cancer Using Multiple Instance Learning in a Multi-task Framework.............................................. 152  
Cheng Li, Sunil Gupta, Santu Rana, Wei Luo, Svetla Venkatesh,  
David Ashely, and Dinh Phung

Shot Boundary Detection Using Multi-instance Incremental and Decremental One-Class Support Vector Machine ................. 165  
Hanhe Lin, Jeremiah D. Deng, and Brendon J. Woodford

Will I Win Your Favor? Predicting the Success of Altruistic Requests........ 177  
Hsun-Ping Hsieh, Rui Yan, and Cheng-Te Li

Feature Extraction and Pattern Mining

Unsupervised and Semi-supervised Dimensionality Reduction and Graph Similarity Constraints............................................. 191  
Zhiyang Xiang, Zhu Xiao, Yourong Huang, Dong Wang, Bin Fu,  
and Wenjie Chen

Cross-View Feature Hashing for Image Retrieval.............................. 203  
Wei Wu, Bin Li, Ling Chen, and Chengqi Zhang

Towards Automatic Generation of Metafeatures.............................. 215  
Fábio Pinto, Carlos Soares, and João Mendes-Moreira

Hash Learning with Convolutional Neural Networks for Semantic Based Image Retrieval............................................... 227  
Jinma Guo, Shifeng Zhang, and Jianmin Li

Bayesian Group Feature Selection for Support Vector Learning Machines... 239  
Changde Du, Changying Du, Shandian Zhe, Ali Luo, Qing He,  
and Guoping Long

Active Distance-Based Clustering Using K-Medoids............................ 253  
Amin Aghaee, Mehrdad Ghadiri, and Mahdieh Soleymani Baghshah

Analyzing similarities of Datasets Using a Pattern Set Kernel............ 265  
A. Ibrahim, P.S. Sastry, and Shivakumar Sastry

Significant Pattern Mining with Confounding Variables..................... 277  
Aika Terada, David duVerle, and Koji Tsuda
### Graph and Network Data

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leveraging Emotional Consistency for Semi-supervised Sentiment Classification</td>
<td>369</td>
</tr>
<tr>
<td>The Effect on Accuracy of Tweet Sample Size for Hashtag Segmentation Dictionary Construction</td>
<td>382</td>
</tr>
<tr>
<td>Social Identity Link Across Incomplete Social Information Sources Using Anchor Link Expansion</td>
<td>395</td>
</tr>
<tr>
<td>Discovering the Network Backbone from Traffic Activity Data.</td>
<td>409</td>
</tr>
<tr>
<td>A Fast and Complete Enumeration of Pseudo-Cliques for Large Graphs.</td>
<td>423</td>
</tr>
<tr>
<td>Incorporating Heterogeneous Information for Mashup Discovery with Consistent Regularization</td>
<td>436</td>
</tr>
<tr>
<td>Link Prediction in Schema-Rich Heterogeneous Information Network.</td>
<td>449</td>
</tr>
</tbody>
</table>
FastStep: Scalable Boolean Matrix Decomposition .................................................. 461
  Miguel Araujo, Pedro Ribeiro, and Christos Faloutsos

Applications

An Expert-in-the-loop Paradigm for Learning Medical Image Grouping ............ 477
  Xuan Guo, Qi Yu, Rui Li, Cecilia Ovesdotter Alm, Cara Calvelli,
  Pengcheng Shi, and Anne Haake

Predicting Post-operative Visual Acuity for LASIK Surgeries ....................... 489
  Manish Gupta, Prashant Gupta, Pravin K. Vaddavalli, and Asra Fatima

LBMF: Log-Bilinear Matrix Factorization for Recommender Systems .......... 502
  Yunhui Guo, Xin Wang, and Congfu Xu

An Empirical Study on Hybrid Recommender System
  with Implicit Feedback ................................................................. 514
  Sunhwan Lee, Anca Chandra, and Divyesh Jadav

Who Will Be Affected by Supermarket Health Programs? Tracking
  Customer Behavior Changes via Preference Modeling ................................. 527
  Ling Luo, Bin Li, Shlomo Berkovsky, Irena Koprinska, and Fang Chen

TrafficWatch: Real-Time Traffic Incident Detection and Monitoring
  Using Social Media ................................................................. 540
  Hoang Nguyen, Wei Liu, Paul Rivera, and Fang Chen

Automated Setting of Bus Schedule Coverage Using Unsupervised
  Machine Learning ........................................................................ 552
  Jihed Khiari, Luis Moreira-Matias, Vitor Cerqueira, and Oded Cats

Effective Local Metric Learning for Water Pipe Assessment ..................... 565
  Mojgan Ghanavati, Raymond K. Wong, Fang Chen, Yang Wang,
  and Simon Fong

Classification with Quantification for Air Quality Monitoring .................. 578
  Sanad Al-Maskari, Eve Bélisle, Xue Li, Sébastien Le Digabel,
  Amin Nawahda, and Jiang Zhong

Predicting Unknown Interactions Between Known Drugs and Targets
  via Matrix Completion .................................................................. 591
  Qing Liao, Naiyang Guan, Chengkun Wu, and Qian Zhang

Author Index ................................................................. 605
Advances in Knowledge Discovery and Data Mining
20th Pacific-Asia Conference, PAKDD 2016, Auckland, New Zealand, April 19-22, 2016, Proceedings, Part II
2016, XXIV, 572 p. 156 illus., Softcover
ISBN: 978-3-319-31749-6